Personality and Ideational Correlates of Student Teachers' Classroom Behaviours.

Thesis

THE OPEN UNIVERSITY

Faculty of Educational Studies
(The Psychology of Education)

Personality and Ideational Correlates of Student Teachers' Classroom Behaviours

by

Alan Butterworth, M.Ed. (Manchester)

A dissertation submitted in accordance with the regulations for the degree of

BACHELOR OF PHILOSOPHY

December, 1976.

Date of submission: 8.12.76
Date of award : 18.1.78

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Statutory Declaration

None of the work reported in this Dissertation has been submitted for another award in this or any other University or Institution.

December, 1976.

Acknowledgements

The author is indebted to those student teachers who consented to take part in the study; to the staff of the Faculty of Mathematics (Student Computing Service), for technical assistance and access to computing facilities; and to Mr. L. J. Chapman, Sub-Dean in the Faculty of Educational Studies, for advice and encouragement during all stages of the work.

A B S T R A C T

Twenty-five second-year students following the Certificate of Education course in a constituent college of the Liverpool University School of Education were tested with personality and creativity measures. The predictive capacity of the personal data was analysed in terms of the several categories of an interaction analysis schedule applied during observations of the students' teaching. The analysis revealed a quantitatively consistent style of interaction when students were seen on different occasions although important changes of a qualitative nature may have occurred. Measures of Ideational Fluency, Flexibility and Originality correlated with the criteria in the same directions and to similar extents, suggesting substantial communality of variance; and there was a tendency for these variables to relate to higher incidences of teacher initiation. Other correlates included Affectothymia (warm, outgoing temperament), artlessness and higher superego strength with the teachers' giving praise; and there were indications that anxiety was linked with teachers' questioning behaviour. Intelligence, as measured by a sub-scale of the personality questionnaire, was not a significant correlate of any behavioural category. The concepts 'personality' and 'creativity', together with their measures, are discussed and related to previous research into teacher competence and behaviour.

CONTENTS

TABLES

TABLES (continued)

FIGURES

CHAPTER 1

STATEMENT OF THE PROBLEM

1.1 The substance of this investigation concerns certain aspects of teachers' behaviours in interaction with pupils and how these may be related to personality characteristics and ideational capacities of the teacher.

Morrison and McIntyre (1973) report that much of the traditional research on teachers and teaching has arisen from practical interest in finding better methods for selecting people who would make 'good' teachers, and in improving the training and assessment of students and practitioners. Such an interest in teacher characteristics and a belief that these bear predictive relationships to the quality of teaching is revealed by the volume of literature, published over several decades, reporting studies which have related personality traits to teacher competence assessed by tests of academic and professional knowledge, and by judgements of practical teaching ability.

1.2 The results of such studies have shown little consensus: no consistent pattern of traits has emerged which could suggest a 'best type' of teaching personality. It is felt that the approach in these researches has been too restricted in its range of predictors, too gross in its applied techniques and also that a lack of carefully defined criteria, appropriate to teaching situations, and of objective assessments of these, have added uncertainty to the findings. This lack of clear pointers to successful teaching may be seen to be attributable to several major sources; two of which are the highly subjective nature of judging the teaching process, and the use of a variety of measures

of personal characteristics whose rationales differ widely and whose outcomes are not comparable except in loosely descriptive terms.

1.3 Several researchers have narrowed down the criterion of competence by separating it from considerations of students' academic performance and have tried to find relationships between practical teaching ability and personality, while others have taken to recording actions believed to have pedagogical significance whilst avoiding, or at least deferring, the evaluation of these as examples of 'good' or 'bad' practice. This present small scale work seeks clarification of the problem of relating personal characteristics to behavioural acts without applying these evaluative criteria which form one of the major sources of difficulty. The main concern of the author is to determine whether specified categories of personal data bear consistently upon observable behaviour, for, in the absence of reliable links between variables of this order, no precise statements can be made about personality and quality of teaching.

CHAPTER 2

REVIEW OF LITERATURE

Criteria of Teacher Success

2.1 Judgements

Probably no area of education has attracted as much research activity as that concerned with describing the attributes of successful teachers, especially those teachers in training or newly embarked upon their careers. As long ago as 1950 Domas and Tiedeman produced an annotated bibliography containing over one thousand references to studies of 'teacher competence'. Most of these investigations and those that have been made since have added to the pool of inconclusiveness by trying to relate teacher qualities to effectiveness of performance - a form of research which has been bedevilled by criterion difficulties in both the 'predictor' and 'dependent' variables.

From a survey of evidence, Evans (1959) finds

> "The most obvious deduction...is that there is no one pattern of successful teacher. Teaching is a complex process calling for many different abilities, no one of which is by itself sufficient to ensure success." (p.33)

The problems of researching in this field have been outlined by Barr (1958), in particular, the formulation of an adequate definition of teaching looms large with the attendant difficulties of specifying criteria as have been discussed by Getzels and Jackson (1963) and Biddle and Ellona (1964).

2.2 Teaching is a polymorphic concept with associated roles and sub-roles. Hoyle (1969) defines two basic sets of roles for the teacher to fulfil in the classroom; one set corresponding with the major functions of instruction, socialization, and evaluation,

and the other concerned with motivating pupils, maintaining control, and generally creating an environment for learning, that is, what can be called facilitating roles. These can be further differentiated into sub-roles of which Hoyle lists fourteen comprising "all the more important ones" (p.59). These can all be regarded as potential sources for criteria of effectiveness in teaching, although some criteria applied in competence studies do not originate in the classroom, especially when assessment takes place within a training paradigm. As might be expected, the researchers who dominate this scene have interests in teacher education and look for criteria of success in their student teachers which take account of a far wider range of variables than those directly observable in the classroom. Researches of this kind typically include indices of proficiency in academic study and are exemplified in Warburton, Butcher and Forrest (1963) who specify eighteen criteria of success, six of which are written examinations, while the work of Cortis (1968) is concerned almost exclusively with examination grades; only one of the six criteria he applies relates to practical teaching.

2.3 The criteria of performance that have been used in many studies are closely concerned with the degree of success with which the college or department courses are negotiated and do not necessarily relate in any major way to assessments of the quality of skill in teaching. This is consonant with the view expressed by Stones and Morris (1972a) who report that many investigations make use of criteria that few would accept as other than trivial and which in many cases are not related to

the act of teaching at all. In any consideration of criteria it is therefore necessary to distinguish those that do bear directly upon teaching acts and those that do not, that is, what we can regard as global judgements which refer to the samplings made of those processes assumed to be important in becoming a teacher, and specific judgements about the competence of pedagogic acts.

The global criteria are not the concern of this present study which deals with the specifics of practical teaching yet the establishing of criteria which will be stable and appropriate for this more limited purpose presents considerable difficulties. In an attempt to sample the bases of judgement in use in colleges and departments of education, Stones and Morris (1972b) posed the question: "How can you tell a good teacher when you see one?" They received completed questionnaires from one hundred and twenty-two institutions in which a final teaching grade of some kind was awarded. Of these, sixty-nine used impressionistic methods of assessing, seventeen used a combination of impressionistic and analytic, only seventeen reported using analytic methods alone, and nineteen did not reply to this question. Almost all the institutions based their assessments upon a series of observed lessons.

However, the popularity of impressionistic methods does not necessarily imply a simple explanation in terms of ease of execution, with attendant suspicions of invalidity. Such methods may be useful, especially when several judges observe the same events but low correlations among supervisors' rankings of attributes assumed to relate to success in practical teaching were found by Robertson (1957). On the other hand, this study

also produced evidence for supervisors' consistency of judgement as individuals. Two related premises for the supervisory system are discussed by Stones (1976): the belief that the lessons observed and commented upon are a reasonable sample of the students' teaching repertoire and the assumption that teaching situations are comparable across different schools. The first of these is clearly independent of competence criteria but the evidence reported by Shipman (1967) is a strong indication that observed lessons are likely to be sufficiently different from the totality of a student's teaching to be an unrepresentative sample of his teaching capabilities. These sampling and 'impression management' effects may be seen as unavoidable aspects of the apprentice-supervisor system which applies universally in teacher education. On the comparability question, it appears that judgements of competence are tempered by considerations of the 'difficulty' of the teaching situation although this is admittedly highly subjective and enters into the grading of performance to varying degrees among training institutions (Stones and Morris 1972a).

The survey by these authors produced no clear pattern of criteria in the questionnaire replies and a factor analysis of the data gave little evidence of conceptual unity. In sum the methods were institutionally and probably personally idiosyncratic. These outcomes are of great importance in that the survey obtained information from a large sample of the U.K. teacher training establishments and they highlight the insecure bases for comparisons of studies which have tried to describe correlates of 'good teaching'.

2.4 There are in existence several publish observational schedules for rating teaching competence and some of these have achieved wide use, especially in the U.S.A. Norris (1975) examined eight American and British schedules which he regards as necessary alternatives to the subjective impressions of individual observers. The application of a schedule permits the building up of a teaching profile through ratings on criteria which remain at a fairly high level of generality. From an inspection Norris grouped the criteria in six main clusters

	% occurrence
(a) Personal and professional qualities	25.4
(b) Lesson preparation	17.7
(c) Statement of aims	6.5
(d) Evaluation and assessment	8.6
(e) Classroom performance	34.9
(f) Children's performance	6.9

(adapted from Norris (1975) pp.88-91)

As Webster (1976) comments,

> "Perhaps the most striking feature of all these schedules (and of impressionistic approaches as well) is the emphasis placed on 'teacher performance' rather than on the amount the children have learnt." (p.81)

This can be partly explained by the practical difficulties posed by the appraisal of pupil gain but the problem of specifying what should be learned is of a philosophical nature. Start (1974) proposed that children's learning should form the criterion of teaching effectiveness and he recommends that it should not be narrowly defined as cognitive nor limited to affective, social or vocational skills. Such a proposition, although logically sound, is fraught with seemingly unsurmountable difficulties since, if observers cannot agree upon what is good praxis in an individual teacher, it is less likely that 'society' would agree

upon educational objectives specific enough to form the bases of criteria of pupil gain yet sufficiently comprehensive to cover all learning that is worthwhile.

2.5 The situation then remains as described by Stones and Morris and provides a background against which to evaluate the studies purporting to investigate the relationships between teachers' personality variables and teaching success which are reviewed in a later section. According to Lantz (1967), there is little evidence that supervisors can say exactly what they are looking for in a student's teaching whether impressionistic or schedule methods are used, and in the absence of firm criteria here, the relationships claimed between personal attributes and competence in teachers remain highly suspect.

2.6 Systematic classroom observation

Many whose professional tasks include making decisions about the relative success of the classroom performance of teachers in training would readily assent to the notion that they can recognise good and bad, or effective and ineffective teaching when they see it. It is less likely, however, that a group of such people would similarly agree upon the criteria by which effective teaching should be assessed.

> "The question: 'What characterises effective teaching?' remains one to which only very inadequate answers can be given. And, as non-significant and apparently inconsistent research results have accumulated, researchers have gradually come to realize that we should not expect much success so long as we seek widely generalisable answers to this question. There is, of course, nothing very original about this realisation, since every practising teacher demonstrates many times each day his implicit knowledge that to have any chance of teaching effectively one must be prepared to change one's behaviour according to the subject matter one is dealing with, the previously acquired

knowledge, skills, attitudes and experience of one's pupils, the size and mood of one's class, and what one is hoping to achieve at any particular moment."

MacLeod et al (1975) p.198

In a final analysis, the effectiveness of teaching must be a function of variables observable in the learner: what knowledge, skills and attitudes have been acquired; but such product variables are less readily accessible to a casual observer or even a more deeply probing researcher. Consequently, much of the earlier research into teaching effectiveness has tended to concentrate upon presage and context variables to the neglect of systematic observation of what happens in classrooms, and what are the results of teaching, i.e. the process and product variables.

Some clarification of these different categories of variables is appropriate and that given by Wragg et al (1975) is as follows:

Presage variables:	those already present when classroom learning commences; characteristics of teachers and learners.
Process variables:	those that describe actual behaviour in the classroom.
Context variables:	descriptors of the classroom environment such as open plan, family grouping.
Product variables:	learning outcomes; acquisition of skills, knowledge, attitudes.

2.7 The focus upon presage and context variables, which is evident in the voluminous research into teacher effectiveness reviewed in the last section, produced few findings of general applicability. There was little consensus on what was relevant to consider and even less on how the chosen variables related

to the diffuse criterion of good teaching. As a result of this and the general neglect of process variables in earlier work, the last decade has seen an upsurge in research into classroom processes. Delamont (1975) reports that classroom research is booming whereas, five years ago, studies undertaken in the teaching situation formed a very minor part of educational research in Britain. A comprehensive description of studies in classroom behaviour is no longer possible in a single book. Simon and Boyer (1970) 'Mirrors for Behaviour' takes up seventeen volumes and Wragg et al (1975) record that British research before 1974 consists of quite a number of reports.

In its early stages this activity in the United Kingdom has been greatly under the influence of the social-psychological approaches exemplified by American workers like Anderson, Bales, Lewin, Lippit and White, and described by Amidon and Hough (1967). Researches in this vein are characterised by a process of coding events into pre-determined categories, and then using the frequency and patterning of codings to place teachers along a stylistic vector according to the constraints they place upon pupils' liberty of expression. Many of the early British investigations of this kind have been influenced by the methods of data collection developed by Flanders and his co-workers as most fully described in Flanders (1970). This style of observing and recording classroom phenomena, called interaction analysis, can be seen to lie towards one end of a continuum which may be represented:

low inference ——————————— high inference

By high inference is meant a system in which reliance is placed upon the observer's interpretation of what is happening, whereas low inference denotes objectivity in the recording of preselected behaviours.

2.8 The majority of research to date reveals a preferred style lying to the left of this continuum although a recent survey (Chanan and Delamont, 1975) indicates some departures towards more phenomenological approaches having roots in sociology and anthropology, and the emergence of a new interest in the structure of discourse in classroom interaction. Eggleston et al (1975) have made the first British attempt to construct a taxonomy of such observational strategies and they outline three major problem areas which affect all interaction studies: (a) what to observe, (b) how frequently to observe, and (c) under what circumstances to observe. Deciding the criteria of relevance under (a) may be seen to present the greatest obstacle to a satisfactory resolution of the difficulties of interpreting phenomena. Philosophers of science such as Karl Popper (Eggleston et al, 1975) make the assertion that there is no such thing as theory-free observation. A major philosophical excursion would be inappropriate here but decisions about what is or is not worthy of attention in any situation are determined by some theory, whether explicit and articulate or intuitive and naive and, in the absence of an established theory of pedagogy, there is a tendency for an observer to attend selectively to events which relate to his own 'theory'.

Underlying all such theories applied in the classroom there are fundamental assumptions about what education is and what it is for, and ultimately speculations of the nature of man: whether

human behaviour is regulated through essentially cognitive channels via information feedback, e.g. (Annett, 1969) or whether our actions are directed by schedules of reinforcement (Skinner, 1968).

2.9 An apparent way of overcoming the problem of relevance of classroom acts is to ask observers to record particular classes of events when they occur but, although this reduces the task to one of recognising instances of pre-determined categories of behaviour, it does not rule out subjectivity or theoretical bias which are then built into the system rather than carried around with the observer and applied in vivo. The use of such category systems can be helpful in that they differentiate the role of observer from that of judge and, while the reliability of the coding is thereby enhanced, the procedure is no less value-free than the methods which range towards the high inference end of the observational continuum. Similarly, the open-minded pre-ordained axis described by Wragg et al (1975) can only be valid at an observational and not at an interpretational level.

Compared with theoretical difficulties of this magnitude the mechanics of observing assume a relatively lower importance but decisions in this area are to a large extent governed by the researcher's approach to (a). For example, Poppleton (1975) describes classroom settings as essentially molar in character in that they are based upon sequences 'which have coherence arising from a unity of task and purpose' (p.251), and argues a case for a more holistic approach. The importance of the ways in which the participants in interaction negotiate their roles is emphasized by Hargreaves (1972) who is critical of interaction analysis which, he believes, can reveal little about the meaning

of events to those involved. These authors favour methodologies which accord significance to meanings and role relationships and these would necessitate consideration of larger units or chunks of interaction.

2.10 The work of Sinclair and Coulthard (1974) in the linguistics field also supports the view that bigger observational units are required. Their studies of transcripts of tape recorded transactions between teacher and pupil, as well as lesson-structuring remarks used to summarize learning or outline future action, frequently revealed a sandwich structure in which a teacher's question was followed by a pupil's reply and a teacher reaction. This triadic sequence of elicit-response-evaluation (Stubbs, 1975) requires larger observational units than conventional interaction analysis, i.e. Flanders', provides. The rationale for this claim is not that the FIAC cannot record sequences of acts but that the procedure precludes the study of how utterances fit together into organised sequences at the level of discourse itself. This view reflects an aspect of the interdependence of the problems posed under (a) and (b).

Eggleston's third problem area concerns the general applicability of any observational system, that is, to what extent this is affected by context variables. Flanders (1973), writing on the comparability of different interaction studies, suggests that apart from the system itself, school setting; the subject matter being taught; the samples of teachers and pupils; and conceptual definitions of the predictor and outcome variables need to be considered.

2.11 The effects of lesson content have been studied and commented upon in this country by Wragg (1972), Morrison (1973) and MacLeod et al (1975). Using Flanders' system Wragg collected data from observations of 102 post-graduate student teachers in secondary schools. A comparison of curriculum subjects showed highly significant differences among the distributions of teachers' verbal behaviours such as their acceptance and use of pupils' contributions, questioning, lecturing (expounding; telling), and giving directions or commands. Differences were also found between the pupils' verbal behaviours 'response to teacher' (answering) and 'initiation', and the proportion of time taken up by silence or confusion. Geography and History showed more lecturing and less pupil initiation. Foreign languages and English showed high levels of pupil talk although there were differences here in that English had a high incidence of both respondent and initiation behaviours whereas few verbal initiatives were taken in the Foreign Language lessons. Other indications of subject-specific variance were found by MacLeod et al (1975) among 33 undergraduate students reading History or English concurrently with Education. Using a research design based upon strategies of questioning and reacting they found that relationships among the 'independent' variables, i.e. the teaching style, varied significantly with the subject content of twenty-minute lessons given to 'micro-classes' of eleven-year-olds. However, in Morrison's (1973) study broad differences in teaching style on FIAC were reported among 25 teachers following the same International Affairs syllabus. This seems to suggest that between-teacher and/or between-class variables may also be potent in controlling interaction.

2.12 The age group of the class was found by Wragg to be related to the distribution of teacher talk irrespective of lesson content. A noticeable and regular increase in the frequency of lecturing to older pupils was accompanied by decreases in teacher criticism and pupil talk; the watersheds occurring between the second and third, and between the fourth and fifth years in secondary schools. Whether this is a true age effect or whether it emerges partly as a result of the increasingly selective function of the secondary school as pupil interests and abilities differentially allocate them to areas of study is not clear, but this outcome serves as a reminder that the age of the 'audience' is likely to be influential in generating distinctive patterns of teachers' verbal behaviour.

2.13 The organisation of the classroom on formal or informal lines is seen by Adelman and Walker (1975) as a factor influencing the efficacy of the observational system. Walker's experience leads him to conclude that instruments like Flanders' can be useful for roughly locating the differences among centralised, formal classrooms but are virtually unusable in informal contexts. The problems of what and how to observe lie at the root of this opinion which arises from an interest in comparing classrooms as entities and not in separate comparisons of teachers' activities. This standpoint is reinforced by the authors' further comment that even with centralised, formal classrooms, what is not collected by the instrument is 'often more interesting and seemingly more representative of the essential quality of the classroom' and 'the decentralised, informal classroom has cultural roles inaccessible to the Flanders' system' (p.221). Rosenshine and Furst (1973) similarly draw attention to the totality

of the education enterprise, commenting that current observational instruments disregard the materials being read, the assignments students write, the teachers' use of written and oral material, the physical features of the room, such as seating arrangements and lighting. There is little doubt that there is an implicit assumption in the naming and distribution of coding categories in systems like Flanders' that class instruction is the norm of teaching behaviour and that verbal interchange is the most salient feature of it. The extent to which such a system can find useful application in less formally organised learning situations is a point worthy of consideration.

2.14 Much has been written by the protagonists of the various approaches to classroom observing and, largely because it was a front-runner in the field and a system with a fairly explicit behaviourist tang, Flanders' Interaction Analysis has provoked considerable comment as is indicated, for example, by references to it in almost every paper published in Chanan and Delamont's 'Frontiers of Classroom Research' (1975). The design of the present work embodies Flanders' system and a preliminary discussion of its origins, structure and applications is pertinent here.

2.15 Flanders (1970) lists ten categories of verbal behaviour that can be observed between teachers and pupils in the classroom: seven are used when the teacher is talking, two when any pupil is talking and the remaining one accommodates periods of silence or confusion. Episodes of observation are divided into three-second segments, for each of which a tally is recorded in the appropriate category. The interaction analysis categories are shown in Table 1. A major feature of the category system is its classification of verbal behaviour in terms of initiatives and responses which are characteristic of interaction between two or more individuals. In this context, Flanders see initiation as making the

Table 1

Flanders' Interaction Analysis Categories

Teacher Talk	Response	1. Accepts feeling. Accepts and clarifies an attitude or the feeling tone of a pupil in a non-threatening manner. Feelings may be positive or negative. Predicting and recalling feelings are included.
		2. Praises or encourages. Praises or encourages pupil action or behaviour. Jokes that release tension, but not at the expense of another individual; nodding head, or saying "Um hm?" or "go on" are included.
		3. Accepts or uses ideas of pupils. Clarifying, building, or developing ideas suggested by a pupil. Teacher extensions of pupil ideas are included but as the teacher brings more of his own ideas into play, shift to category 5.
		4. Asks questions. Asking a question about content or procedure, based on teacher ideas, with the intent that a pupil will answer.
		5. Lecturing. Giving facts or opinions about content or procedures; expressing his own ideas, giving his own explanation, or citing an authority other than a pupil.
	Initiation	6. Giving directions. Directions, commands, or orders to which a pupil is expected to comply.
		7. Criticizing or justifying authority. Statements intended to change pupil behaviour from non-acceptable to acceptable pattern, bawling someone out; stating why the teacher is doing what he is doing: extreme self-reference.
Pupil Talk	Response	8. Pupil talk - response. Talk by pupils in response to teacher. Teacher initiates the contact or solicits pupil statement or structures the situation. Freedom to express own ideas limited.
	Initiation	9. Pupil talk - initiation. Talk by pupils which they initiate. Expressing own ideas: initiating a new topic; freedom to develop opinions and a line of thought, like asking thoughtful questions; going beyond the existing structure.
Silence		10. Silence or confusion. Pauses, short periods of silence and of confusion in which communication cannot be understood by the observer.

first move, leading, beginning, introducing an idea or concept for the first time, expressing one's own will; and responding is regarded as taking action after an initiation, conforming or complying with the will expressed by others. Thus the preponderance of categories for teacher talk within the system reflects the expectation that the teacher will, in most situations, show more initiative than the pupils. Interaction analysis using this system apparently ignores non-verbal actions but these can be crucial in determining the categorisation of an utterance insofar as they indicate or amplify the speaker's intentions. Morrison and McIntyre (1973) note that the technique preserves a considerable amount of actual behaviour and this opinion lends credence to the common assumption that what is said in a situation is a form of behaviour sufficiently significant to be treated in isolation from the other forms of behaviour without undue distortion in recording the interaction. It can be argued here that verbal communication is the dominant mode of teacher-pupil interchange and by sampling it we can best approach the problem of quantifying total classroom behaviour within the limitations imposed by an observer's capacity to accurately assess the ongoing activity.

2.16 Flanders' system was constructed as part of a study in which teacher influence was related to pupil achievement and attitudes (Flanders, 1965). The hypotheses tested in this process-product investigation were concerned with the effectiveness of direct and indirect teacher influence conceived in terms of the initiation and respondent behaviours mentioned above. This cast teacher behaviour as a predictor variable which needed quantification and it was for this purpose that the system was devised. The teaching behaviours chosen for recording were closely related to those identified by prior research into classroom climate, e.g. Medley and Mitzel's (1963) fairly

high inference categories:

(a) emotional climate – the relative amount of hostility observed

(b) verbal emphasis – the relative emphasis on verbal and traditional schoolroom activities

(c) social structure – the relative degree of pupil-initiated activity

Direct influence subsumed verbal statements by the teacher that restrict freedom of action by focusing attention on a problem, interjecting teaching authority, or both. These embrace the behavioural categories 5, 6 and 7 in fig. 1. Indirect influence was described as verbal statements by the teacher that expand a pupil's freedom of action by encouraging his verbal participation and initiatives. Acts under this heading are recorded in categories 1, 2, 3 and 4. The categories themselves were chosen to be exhaustive of all possible types of verbal interaction within the compass of the system and to allow an observer to distinguish between those acts which resulted in compliance and those which invited more creative and voluntary participation, without his being diverted by the subject matter.

2.17 Thus it may be seen that the system is designed to classify teacher verbal behaviour, and pupil behaviour is important only insofar as it indicates the apparent effects of what the teacher does. Nowhere does its author claim to sample all that is important in a classroom as some critics of the system, already referred to, seem to imply. The emphasis in Flanders' writings (1965) (1970) is upon the outcomes of direct and indirect teacher influence upon pupil variables which provide a rationale for attempts to increase the capacity to exercise indirect teacher

influence in both established and beginning teachers. Evidence indicating that teachers who were able to exert indirect influence increased the achievement of their pupils is difficult to find outside the work of Flanders and his disciples. There have been a few British studies which have detected styles within the FIAC system and on similar instruments but which have not found these to be related to differential achievement in pupils. Morrison's (1973) finding of broad stylistic differences among 25 teachers was not related to differences on tests of knowledge applied to the 400 children they taught. Eggleston's (1973) findings concur. Here 95 science teachers were observed using a 25 category system which permitted them to be allocated to one of three stylistic groupings ranging from 'highly dominating of interaction' to allowing 'pupil-centred enquiring'. Process-product analysis related teaching behaviour to tests of pupil achievement and revealed that most of the variance was accounted for by differences in initial ability rather than teaching style.

In the area of pupil attitudes there are indications of greater accord among the research findings. Flanders (1970) summarises seven projects supervised or directed by him between 1955 and 1967, involving 112 classes, and in which pupil attitudes were among the product variables. Commenting upon the findings, Flanders writes:

> "The same (results were) found in Minnesota and New Zealand, some 8,000 miles apart, in spite of differences in teaching style and pupil expectations. The teachers of classes that scored high on liking the teacher, motivation, fair rewards and punishments, lack of anxiety, and independence used more indirect influence, while teachers of classes that scored low used less indirect influence...The greater use of indirect influence meant asking more questions,

> clarifying and using pupil ideas, and giving praise... One gets the impression that a small amount of indirect influence lubricates the classroom gears of subject matter learning, and even though the total amount of indirect influence is small, its presence or absence (is significantly related to) the positive or negative attitudes of pupils, respectively."
>
> (p.392, parentheses in original)

2.19 There are also indications in, for example, Bellack et al (1966), Wright and Nuthall (1970), Hughes (1973) that the teachers' reactions to pupil contributions of initiatives may be strong predictors of attitude and, indirectly, of achievement. At a commonsense level there seems little doubt that when pupils can experience the acceptance of their own ideas for incorporation into the fabric of the lesson they will learn that what they have to contribute is worthy of consideration by the teacher and a corresponding increase in positive attitudes toward learning and involvement in verbal interchange will ensue. Whether this works via conditioning or information feedback would seem to be an arid argument; that it does seem to occur is important in its implications for teachers and those who teach them.

MacLeod (1975) discusses the operation of indirectness via the teacher's reacting strategies; a reaction being defined as a teacher behaviour elicited but not solicited by a previous pupil contribution. Bellack et al (1966) amplify this definition to include rating (positively or negatively) and/or modifying (by classifying, synthesising or expanding) the pupil behaviour which occasioned them. Thus lessons with a high proportion of verbal interaction can be considered to consist of sequences of dialogue in which pupils' respondent behaviours are elicited in the early

stages and there is a following shift to a flow of respondent and initiation behaviours regulated by the teacher's reactions, these taking the forms of praise, criticism, explorations on the basis of pupils' expressed ideas, and further questioning.

2.20 Interaction analysis systems have been criticized for being concerned with the minutiae of classroom occurrences (Nuthall and Snook, 1973) but the allocation of a single category to teacher questioning behaviour in the Flanders system seems too gross a classification since questions can be asked which make widely varying cognitive demands upon pupils; e.g. one can imagine simple questions designed to elicit factual recall of learned material or questions of a deeply probing nature which function as pivotal points in the dynamic of the lesson. Category 4 of FIAC cannot differentiate between such extremes and the lumping together of such a wide range of behaviours which vary in complexity and function is an apparent weakness of the system. The scoring of n tallies in category 4 can reveal little about what is widely considered to be an important aspect of teacher performance, that is, the strategies of questioning adopted. A teacher's behaviour encoded in this category can represent a mere harping on the same trivial theme or a skilful deployment of questions to probe understanding of subtle relationships, or to direct pupils' attention to as yet unconsidered aspects of a topic. The varied nature and functions of questioning are described by Flanders (1970) in terms similar to those outlined here yet one is left to infer the type of question a teacher has asked from a perusal

of the sequence of scoring tallies in a matrix.

> "The pattern of open questions is illustrated by the 4-9 and 4-8-9-9 transitions. These questions stimulate pupils to express their own ideas and to contribute their own suggestions. This kind of pattern is most effective when seeking to lift the level of abstraction in order to set an issue in a broader context, when the opinions and reactions of pupils are solicited during the planning of work, and when speculation and explanation is to come from the pupils."
>
> (p. 283-4)

Even in a refined form, where sequential tallying of categories is intended to provide useful information on the progress and direction of the interaction, the system remains crude in its treatment of questioning behaviour. But the problem of classifying questions remains outside interaction analysis and is not a product of it; particular difficulties are met in comparing studies which group questions according to their cognitive demands with those which are concerned with the scope of pupils' freedom in choosing the kind of response they make (MacLeod, 1975).

2.21 The 'open-closed' typing of questions implied in the above extract provides an example of the dichotomous classification which Rosenshine and Furst (1971) consider to have been a hindrance in providing consistent and meaningful conclusions about the nature of effective questioning, despite findings of several significant relationships between teachers' questioning and pupil outcomes. On the other hand, the generation of more categories seems unlikely, of itself, to be profitable (Rosenshine and Furst, 1973); the clarity and explicitness with which the dimensions underlying the categories are conceptualized is more

fundamental. MacLeod (1975) refers to the semantic confusion which abounds:

> "One can never be sure whether the 'factual' category used by one investigator is conceptually related to the category of 'closed' used by another, or whether one man's 'open' question is another man's 'divergent' question."
> (p. 203)

2.22 Whilst criticizing Flanders' system for its failure to cater for different 'levels' of questioning it is salutary to bear in mind that observational schedules which have attempted to do so have borne little fruit and simply to note, at this stage, that category 4 tallies indicate undifferentiated questions. This aspect of the system would prove a serious impediment in a study which had questioning strategies as part of its concern and, in that case, the use of FIAC would be contraindicated. In studies where the focus of interest lies in gross verbal interaction or the relative distributions of wider pedagogical acts the Flanders System is more appropriate.

2.23 Interaction analysis has found application in teacher education in two related areas: the training of students in observation of teachers in action; and the provision of feedback on students' own practices in the teaching situation in order to 'improve' style in accordance with some prescription of good technique. The training milieu for such work has often been one of microteaching which is

> "a scaled down but realistic classroom context which offers a helpful setting for a teacher (experienced or inexperienced) to acquire new teaching skills and to refine old ones" (McKnight,1971).

The bringing together of these two aspects of applied interaction analysis has formed the basis of courses of professional training in teaching. Kleinberg (1975) sees the transition taking place

from using interaction analysis to identify, classify and quantify interactive behaviour in classrooms to using components of interaction as the basis for defining skills and building a repertoire of such skills in parts through microteaching experience.

The present concern relates to the effects which knowledge of an observational system has upon a student's interaction style since such knowledge is available in the professional course being followed by those included in the sample in this investigation. (See section 4.6).

2.24 Brown (1975), commenting upon case studies of teacher preparation, reports that students' perceptions of teaching can be changed by training in interaction analysis; and Hough et al (1969) from a study involving 400 students, concluded that their experimental group, having been taught Flanders' system, used significantly more praise and encouragement, more acceptance and clarification of pupil ideas and less criticism than those whose training had included unstructured observation of microteaching. These results indicate clearly that students' acquaintance with an observational system may affect their own performance in a classroom and this is a contingency for which allowance needs to be made in the present work.

2.25 Personality

The subject of this section has a legitimate claim to the status of a general theory of psychology and, as such, cannot be treated here in anything approaching a comprehensive manner. According to Allport (1937) there are at least fifty different

meanings of the term 'personality'. Consequently, the writer's intention is circumscribed to that of merely providing a sufficient theoretical context in which to locate the related aspects of other areas of the literature review, the experimental design, and discussion of findings.

According to Semeonoff (1970):

> "Crude misconceptions apart, Personality, like Intelligence, is a term which probably causes much less trouble to the informed - or even the uninformed - layman than to the psychologist."
> (p. 9)

Such a lay view of personality could perhaps be summed up as that which makes one man different from another. Taking this a step further, it is recognized that although people are not entirely consistent in what they do, how an individual will react in a given situation can be predicted from a knowledge of his or her personality. There is, as Cattell (1965) remarks, observable 'predictive dependability' in the behaviour of those around us, that is, an orderliness in one's sensations and feelings, and a rationality that militates against the notion that personality must be for ever unpredictable. It is common experience that we can predict with considerable accuracy how friends and members of our own families will react when presented with various stimuli and situations.

2.26 This consistency of disposition to actions producing characteristic behaviour forms the fabric of many otherwise diverse theories of personality. Detailed discussion of these is inappropriate but a sufficient indication of their orientation may be had economically by a consideration of classes of definitions

advanced by their authors. Such a taxonomy of definitions has been put forward by Guilford (1959) who lists, inter alia, stimulus, omnibus, integrative, and adjustment types.

2.27 The idea that personality is an individual's social-stimulus value comes close to the popular view that equates personality with social adroitness; the ability to impress others, which is generally stated in terms of a typical single attribute, for example, aggressiveness or persistence, with the implication that one may have 'personality' or not. In similar vein, the term is occasionally used to indicate a capacity to endure hardship or responsibility when it takes on the strongly cultural denotation of 'character'. Other forms of this usage suggest marked individuality or eccentricity as, for example, when used in a context like "He is quite a personality." Definitions of this kind are open to serious criticism but have some merit in that they emphasize that it is only by its effects upon others that personality can be appreciated. A logical extension of this stimulus view is that any object can only be defined by the impressions it makes on its observers and, as a person's effect upon others is not a function of himself alone, he might have as many personalities as there are people who know him. Guilford (op.cit.) concludes that personality so-defined would involve an appraisal of the perceptions, prejudices and inferences of those who know the individual. Vernon (1963) describes the social-stimulus view as a naive interpretation of personality in which we judge people by observing their outer appearances and expressions, how they behave in various contexts,

and what they or others tell us about them. The reciprocal interaction which occurs between any two persons has evaluative components where each judges the other and tries to behave in such a way as to create a favourable impression. The result of this aiming to impress is the creation of mask characteristics (Eysenck, 1947) which are distinct from the substance of personality. It is interesting to note, in passing, the widely reported derivation of the word 'personality' from the Latin persona, meaning a theatrical mask.

2.28 Some attempts to arrive at comprehensive definitions have produced omnibus or 'rag-bag' statements which are now of mainly historical interest. Probably one of the best-known examples is that given by Prince (1924) who refers to the sum total of biological, innate dispositions, impulses, tendencies, appetites and instincts of the individual and the acquired dispositions and tendencies. Such a lumping together of concepts yields a mere aggregate of qualities that does little to promote systematic study of the topic, whereas an integrative type of definition emphasizes the organization of personality, for example,

> "....the integrated organization of all the cognitive, affective, conative and physical characteristics of an individual as it manifests itself in focal distinctness from others." (Warren, 1934)

This type of statement stresses functional unity and also clearly highlights the uniqueness of the individual, a feature that finds acceptance in many theories, especially those constructed from an idiographic perspective, and which act as a kind of corrective to anthropological views (Marcuse, 1955).

2.29 Several theorists include concepts of adjustment in their statements which is consistent with a behaviourist view of

psychology as the manifestations of an organism's coming to terms with the environment. By this line of thinking personality becomes an individual's unique pattern of adjustments. Eysenck (1952) sums up the view expressed in much American psychological literature by remarking that there are no general and consistent forms of conduct, which, if they did exist, would make for consistency of behaviour, and stability of personality, but only specific S-R bonds or habits. Personality is thus placed firmly in the behaviourist tradition which regards it as a composite of specific habits acquired through learning. However, the concern with objectivity and measurability which is usually thought to be the hallmark of this school of thought is sometimes tempered by evaluative criteria as in Watson's (1925) view of personality as an individual's total assets and liabilities, actual and potential, on the reaction. Guilford (op.cit) concludes that if all behaviour is adjustment it might as well be said that personality is an individual's characteristic pattern of behaviour, but to introduce a limitation of meaning to social adjustment, as some definitions do, would have only minority support among most of those who deal with personality. (p.5)

Implicit in much of this treatment of the subject is the latent controversy of the specificity or generality of personality; whether or not it is justifiable to regard aspects of individual qualities and behaviours as quantitatively different indications of the same dispositions which are common to all. The specificity view has been briefly stated in Eysenck's summary already referred to; at the other extreme is the belief that there are general tendencies which, throughout life affect all human activity

regardless of the type of situation in which they occur. Jung's (1933) theory of orientations towards introversion and extraversion provides an example. As with many similar academic debates the truth probably lies between these poles. The classic research into deceitfulness and other traits in children by Hartshorne and May (1928, 1929) suggests this solution in which behaviour is neither completely dependent upon the situation nor general and independent of it. The view expressed by Allport (1966) places the emphasis clearly on the individual's side of this range. Situations and sociocultural variables may be distal causes but

> "...the intervening factor of personality is ever the proximal cause of human conduct... It is the person who accepts, rejects or remains uninfluenced by the social system." (p.9)

Cattell (1965) relates the context of behaviour and the role it requires the individual to play and he opines that a lack of allowance for the situation is one of the main causes of misjudging personality. The same author also clarifies the relationship between the concepts of role and personality in which he holds the latter to be superordinate.

> "Role is part of personality in the broader sense, for to react in a role the person has to acquire what it takes to do so. (p.27-28)

2.30 If it is accepted that the specificity-generality issue is one of emphasis rather than absolutes what is the nature of the more specific and the more general qualities of personality and what would be a sufficiently comprehensive and yet precise definition to accommodate them in a systematic way? Taking the last question first, the statements by Allport (1937, 1961) may be considered to combine the most refined elements of types of

definitions, some of which are reviewed here, while avoiding their shortcomings.

> "Personality is the dynamic organization within the individual of those psychophysical systems that determine his...unique adjustments to his environment (1937)
>
> ....characteristic behavior and thought (1961)

This definition in both its forms can be seen to be a scholarly one and truly eclectic in its integration of those constituents of earlier statements which have found most general acceptance. In it personality is an organisation, that is, there are systems that relate the components and allow the possibility there will be some disorganisation. It is also noteworthy that the organisation is dynamic as opposed to static: personality develops and changes, with important implications for those who would claim to measure it. Furthermore, the emphasis is upon inner aspects, <u>within</u> the individual, and this places the statement in the category of substance definitions. The description of the systems as psychophysical reaffirms their disposition to action.

> "(It) reminds the reader that personality is neither exclusively mental nor exclusively neural. The organization entails the operation of both body and mind, inextricably fused into a personal unity."
>
> Allport (1937) p.48

That the organisation of such systems determines adjustment makes the point that personality is not synonymous with behaviour nor the effect one has upon others, but that the systems of dispositions bear a causal relationship to the mask aspects which are observable. The claimed uniqueness of the adjustments indicate the importance Allport's theory ascribes to individuality: every adjustment being unique in time and place generates a vast

diversity of experience; no two individuals will have the same personality, although there will be sufficient similarities among personal qualities which can be considered as common. It is these dimensions that Allport designates as traits.

2.31 Adjustment has been mentioned but the sense in which it is used in this definition does not seem to mean mere reactive adaptation but includes spontaneous behaviour in which persons shape their environment as well as being moulded by it. Hall and Lindzey (1970) discuss this in connexion with Allport's (1961) amendment to his statement and see this shift in emphasis as a reflection of developments in existential thought. This, they report, created a reaction against the traditional view of adjustment with its connotations of 'giving in to' or 'submitting to' outer forces and demands. Not that these demands need always to be outer in origin: it seems possible to accommodate the Lewinian concept of psychological environment within the term.

2.32 Allport's definition has cleared the ground of much conceptual lumber and, in particular, has established the idea of personality traits which has been taken up by subsequent theorists and applied in several senses. Problems of using trait descriptions centre on three main issues: the relative status of different traits, the methods by which they are derived, and difficulties of causality.

The systematic methods pioneered by Allport employ a combination of these approaches in the labelling of personality dimensions. In collaboration with Odbert (1936) Allport listed almost eighteen thousand English words used to distinguish the behaviour of one person from that of another, and from this material refined 'authentic traits of personality' by rejecting

those words which indicated temporary states or were evaluative terms. The techniques used here were pre-quantitative in that the process was one of careful analysis of the meanings and contexts of application of the personality descriptors, in contrast with the identification of traits by cluster or factor analysis by, for example, Cattell, whose work had a similar starting point in word lists but whose products assume the name 'traits' in a different sense, that of more fundamental qualities which do not directly correspond to behaviour observable and classifiable in terms of the original words. The identification and naming of traits is a three-step process according to Carr and Kingsbury (1938). It firstly involves the observation of certain adverbial characteristics of the individual's behaviour, then these qualities are transferred from the action to the actor where they come to describe the person whose actions are affected by them. The reactive nature of the individual is then described in adjectival terms, and finally these qualities of the individual, rather than his actions, are abstracted and referred to as things. By this derivation it can be said, for example, that a person has a trait of persistence. This shows that traits are essentially abstractions and do not necessarily refer to structural aspects of the individual. The quantitative approach which searches for underlying dimensions by examining correlated traits introduces a fourth level of processing but both approaches apply the principle of parsimony in tackling the problem of reducing thousands of trait names, many of which must be near

synonyms or opposites, in order to arrive at a smaller number of relatively independent categories or major dimensions for detailed study. It is here that Allport's theory is weak in that it cannot specify a set of dimensions to be used in personality study. By their definition individual traits cannot be stated in a general form and an investigator who takes the idiographic approach is constrained to begin afresh the task of devising variables for each individual he undertakes to study. This is not an encouraging prospect for a researcher.

2.33 The other crucial question regarding traits and dispositions, that is do they guide or direct behaviour only, or do they initiate it, is a very fundamental one on which Allport (1961, 1966) indicates that some traits are more impelling and have a more crucial motivational role than others, but he leaves the issue unresolved, unlike Cattell (1965) who comes out clearly in causally relating 'source traits' to behaviour through 'specification equations'. The approach of this author will be reviewed briefly here for two reasons: firstly, because it can be viewed as an extension of trait theory by the application of quantitative techniques and, secondly, because it is within this theoretical framework that part of the present investigation is conceived.

2.34 Cattell (1950) gives a very general definition of personality as "...that which permits a prediction of what a person will do in a given situation." (p.2) This statement does not provide an axiom from which the theory springs but is an interim formulation pending a complete description of the concepts to be employed in behaviour study as the theory continues actively to develop and change. However, the emphasis upon prediction in it indicates

the orientation to be 'scientific' and this is borne out by the sophisticated multivariate factorial methods which characterise the research designed to map out the salient features of personality structure. Cattell is widely regarded as the principal exponent of the quantitative and objective approach to the study of personality and, according to Wright and Taylor (1970), he has the "....reasoned conviction" that the only valid theory is one which begins with, and is based upon such techniques, and that he "emphatically rejects any approach that begins with armchair theorising." (p.534)

The main structural concepts of the theory are traits although this term has significances different from its usage in some earlier trait theories; the main difference being the various levels or orders of traits and their causal as contrasted with their descriptive functions. Cattell makes a distinction between surface and source traits. Surface traits may express themselves directly and may be regarded as clusters or a number of correlated consistent qualities, whereas source traits are revealed by factors and are major causal influences in behaviour. This directness of expression of the surface traits may appeal to a common-sense observer as a more valid and meaningful structure than that afforded by the source traits because it corresponds to the kind of generalisations that can be made on the basis of simple observation, but the source traits, revealed by factor analysis, are claimed to have the most utility in accounting for behaviour. It is only by approaching the problem by multivariate methods that the pitfalls of falsely identifying

surface traits as the basic structural components of personality can be avoided.

> "(Source traits) promise to be the real structural influences underlying personality...as research is now showing, these source traits correspond to real unitary influences - physiological, temperamental factors; degrees of dynamic integration; exposure to social institutions - about which much more can be found out once they are defined."
>
> Cattell (1950, p.27)

2.35 How then are the source and surface traits derived and what are their qualities? Theoretically there might be as many surface traits as polar adjectives which refer to behaviour in a culture. The task is one of sifting and reducing attributes to yield a useful and economic set of dimensions as was done by Allport and Odbert (1936) and in this Cattell and his co-workers have applied the factor analytical methods originally developed during the 1920's by Thurstone in his investigations of cognitive structures. The procedures adopted appear to follow a three-stage process: (i) a thorough survey by examining a wide range of personality variables, (ii) establishment of major dimensions by factor analysis, (iii) validation by comparison of predictions with actual behaviour.

2.36 The survey stage started with the Allport and Odbert list in which synonyms and close terms were brought together to produce a basic vocabulary of 171 terms describing the whole personality sphere (Cattell, 1946) and for which methods of sampling were devised within three modes. Cattell and Butcher (1968) summarize these as observation, self-evaluative or questionnaire data, and objective test data, yielding 'L', 'Q' and 'T' data respectively.

'L-data' consists of actual records (time-sampling counts, behaviour ratings), personality manifestations in everyday life (occupational performance, automobile accidents). 'Q-data' is derived from introspection and replies to set questions, and 'T-data' comes from the individual's behavioural responses in miniature real-life situations.

In a large number of factor analyses of correlations among measures in these modes some sixteen to twenty-one source traits have been isolated. These are otherwise denoted 'primary' or 'group' factors, the best-known of which are those which form the basis of the 16 Personality Factor Questionnaire and which are shown in the following table. (Adapted from Cattell, 1965, p.365)

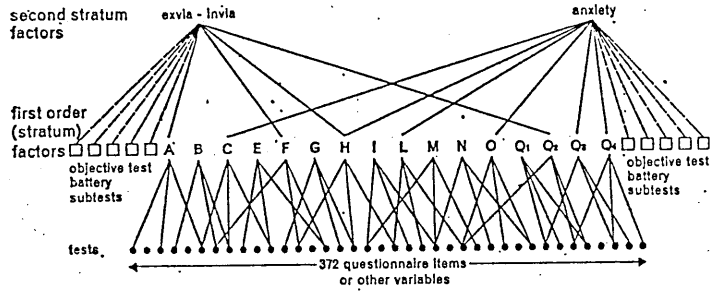
Table 2

Factor Symbol	Technical Label	Popular Label
A	Affectothymia-Sizothymia	Outgoing-Reserved
B	Intelligence	More intelligent - Less intelligent
C	Ego strength	Stable-Emotional
E	Dominance-Submissiveness	Assertive-Humble
F	Surgency-Desurgency	Happy-go-Lucky - Sober
G	Superego strength	Conscientious-Expedient
H	Parmia-Threctia	Venturesome-Shy
I	Premsia-Harria	Tender-minded - Tough-minded
L	Protension-Alaxia	Suspicious-Trusting
M	Autia-Praxernia	Imaginative-Practical
N	Shrewdness-Artlessness	Shrewd-Forthright
O	Guilt-proneness-Assurance	Apprehensive-Placid
Q1	Radicalism-Conservatism	Experimenting-Conservative
Q2	Self-sufficiency-Group adherence	Self-sufficient - Group tied

Q3	High self-concept - Low integration	Controlled-Casual
Q4	Ergic tension	Tense-Relaxed

2.37 These factors are not independent in the mathematical sense and their intercorrelations represent residual variance which is capable of being resolved by further analysis to give second-order factors. Some six of these have been found, of which the two most important have been labelled Exvia-Invia and Anxiety (low-high). The links between the primary and these second-order factors are shown in the following figure, (reproduced from Cattell, 1965, p.118), the linkages indicating either positive or negative relationships.

fig. 1



The second-order factors are seen as 'typical organizers of primary factors' (Cattell, 1965).

> "They are analogues to higher executives in a hierarchy in that they do not have immediate, intimate, effect on the lower operators but act on them indirectly only through their direct influence on the intermediate controllers (the primary factors). (ibid, p.117)

The technical names of the factors are interesting in that Cattell avoided using words which would ordinarily denote surface traits since, by definition and the process of derivation, the source traits cannot be precisely identified with any common-sense label, yet it was found necessary to gloss each neologism to indicate in plain language what the meaning is like.

The last four Q factors are unique to questionnaire data which provide the only source of subjective (i.e. introspective) information for analysis (Cattell and Saunders, 1950). These four apart, there has been some modest agreement among the factor patterns that have emerged from analyses of data from the three modes. That there is less than a good correspondence across data sources may merely indicate (Cattell, 1966) that the different measurement approaches sample data at different levels of generality so that a one-to-one match of factors is not found.

2.38 A cornerstone of validation for this theory is the consistent appearance of dimensions independent of data source when the factors are rotated 'blind' to oblique simple structure. Hall and Lindzey (1970) comment on this that Cattell's initial hope of finding identical factor structures in all three data sources has been realized only partially. In the absence of any firm predictive capacity this would seem to be a serious flaw in the theory but there are three claimed relationships which support the validity of the factors found. Firstly, there seems to be a clear resemblance between some of the factors and dimensions already recognized by other theorists. For example, factor 'A' corresponds to the temperamental dimension schizothymia - cyclothymia described by Kretschmer (1925), and factor 'I',

roughly described as tender-minded vs. tough-minded is anticipated in the writings of William James.

Secondly, the reappearance of the same dimensions in diverse cultures seems to suggest that the theory does deal with basic personality dimensions. This ubiquity of trait configurations is claimed by Cattell and Butcher (1968) to show facets which are characteristic of human nature wherever it appears.

Lastly, predictive validity is claimed from the successful application of the 16 P.F. Questionnaire in a variety of situations including occupational differences and requirements, the prediction of leadership in small groups, the identification of neurotic individuals (Cattell and Butcher, ibid.). This route to validation operates via the 'specification equation' (Cattell, 1965) which involves the prediction of actual performance and behavioural response from scores on individual factors. The simplest way to combine such sources is by summation in the proportions in which they normally contribute to success, or magnitude of response in a given situation, e.g.

$$(B) = 0.5A + 2B + 1.2C - \ldots\ldots$$

In view of the complexity of the theory as a whole the specification equation provides an extremely simple model which implies that each source trait has an independent and additive effect upon the response. Cattell (1956, 1965); Cattell & Butcher (1968) recognize that in some cases the relationship may not be linear and that there might be interaction among the factors, but in the absence of evidence of such interaction it is felt appropriate to proceed with the simplest

kind of estimate, which often provides good approximations to more complex models and because it "will facilitate calculations made by teachers and other users of available measures." (Cattell & Butcher p.59)

2.39 Criticism of Cattell's theory fall into categories which can be termed general and specific. Much of the general comment concerns the nature of trait theories and the vicissitudes of factor analytic methods applied in any context. The positive aspects of trait theories lie in the attempts made to rationalise the conceptual bases of classification to permit parsimonious and useful descriptions to be made. A breaking down of the total personality field can produce more manageable items and make consistencies and generalities readily amenable to study. On the other hand, individual trait scores along cannot provide an adequate description of personality even though it may be an accurate one; there must be some indication of dynamics and hierarchical ordering of these in human activity. Vernon (1963) applies a pragmatic test of the adequacy of trait theories and concludes that they have not worked well enough because the huge volume of research seems to lead into a dead end. He discusses the expectation that trait theory could effect a rapprochement between the idiographic and nomothetic approaches but sees the opposition of these to be as strong as ever.

Another objection to trait theory outlined by Hilgard and Atkinson (1971) is the implication that the assigning of traits to an individual asserts something fundamental about the way he is, causing him to do what he does. Conversely a trait may be considered to be an assertion of a capacity to behave in certain

ways under defined environmental provocations. For example, Bandura and Walters (1963), in their study of aggression, found that boys were not aggressive under all provocations. This last criticism would not be acceptable to a factor theorist who would dismiss it as applying only to surface or descriptive traits as in the studies by Hartshorne and May, but there are several points on which the factor analytical method itself is susceptible to attack.

Lykken (1971), in reviewing the application of factor analysis to personality research, distinguishes two types of use, data reduction and hypothesis-testing which he considers legitimate, and attempts to discover and specify a structure of personality as Cattell has done. On this latter use Lykken is sceptical that the techniques do produce meaningful information out of complex material and argues that factor analysis of 'known' synthetic data might indicate the appropriateness or otherwise of the method. From experience of this kind he concludes that factor analysis does not provide useful information that was not discernible in the original data. Similarly, Brody (1972) who reviews the evidence concerning Cattell's primary factors, finds that a description of personality based upon their use rests upon faith rather than facts. Allport's (1937) criticism is of a more fundamental kind. He alleges that factor theorists create systems of artifacts that have no true relation to any single individual and consequently distort and misrepresent reality. The essence of these objections seems to be that the derived factors are not psychologically meaningful. That they

are artifacts is beyond dispute but the critics are claiming that they do not fit the observations of other students of human behaviour. The counter-argument is that it is the explicit intention of the factorist to go beyond what is simply observable with the expectation that the accumulation of empirical findings will make factors more useful than the variables now in common use.

2.40 Factor analytical methods can only find what has been fed into the computer. In his search for objectivity the factorist may simply be moving subjectivity from its conventional location in the interpretive step and introducing it earlier at the point where he decides what tests or measures to include in his correlation matrix. This criticism is a valid one in those cases where data has been collected over a narrow front of activity but Cattell's sampling of a widely defined personality sphere seems less subject to it. Hall and Lindzey (1970) affirm that even if factor analysis depends upon prior ideas, it provides a means of assessing the fruitfulness of those ideas and this is a constructive view to take of a still-developing theory. They point out

> "In contrast, many personality theorists have originated hosts of personality variables without ever submitting them to the empirical crucible." (p.410)

The recurrent problem of causality needs ever to be borne in mind when assessing factors as springs or results of behaviour. Correlation is not a vector quantity in this sense, consequently questions about the direction of causality can only be answered by designing longitudinal studies in which 'before and after the event' tests can be applied. Cattell and Butcher (1968)

record the urgent need for such studies as extremely little work has been done on this problem.

2.41 Some specific criticism of Cattell's theory stems from the large scale of his approach where efforts to chart the whole domain of personality structure have precluded the detailed analysis of any one portion of the whole task. That Cattell has frequently claimed a surer empirical basis for his constructs than actually exists has formed the substance of criticism by Becker (1960), Vernon (1963) but it seems almost inevitable that large amounts of any exhaustive personality theory should rest on less than firm foundations during the theorist's lifetime.

2.42 On the positive side, the operational nature of the theory with its requirement of empirical activity has ensured an integration of the conceptual framework with experiment which continues with vigour. Personality study is an area of psychology that from its inception has been greatly influenced by subjective approaches and the emphasis upon explicitness and quantitative methods which Cattell's theory places can be seen to be a healthy development. It is considered that the theory provides a set of dimensions appropriate for the present study of those personal attributes which hold a predictive relationship with observable behaviour.

Questionnaire Measures of Personality

2.43 There appears to be little doubt that the questionnaire, by virtue of its ease of application is the most commonly used personality measuring instrument. But, in spite of its popularity, the self-report method is widely claimed to be susceptible to serious distortions which adversely affect the validity. These

distortions may originate in one or more of four ways: deliberate faking; misinterpreting questions; response set; and instrumental factors.

Thorndike (1971) believes the major concern of authors of questionnaires to be the ease with which subjects can deliberately distort their answers to present a personality image they consider to be desirable. These aberrations in an individual's reports of his own typical behaviour would seem to be most marked when there are incentives to create a favourable impression but Vernon (1963) considers them also to occur in survey situations where the subject is guided more by his self-concepts of his traits or attitudes than by the content of the questions as such.

2.44 Cattell (1965) reports that when subjects respond as truthfully as possible in a situation which offers anonymity, and then do the same test again where they can be overtly identified and the results have a bearing on their future, the latter situation does not produce an aberration along a single dimension of social desirability but along several, depending upon the role which the subject chooses to play. If Cattell's view is credible this effect would seem to undermine any attempt to reduce these distortions simply by manipulating the test content without reference to the circumstances under which the test is applied. It does seem probable that motivational distortion can be mitigated by reducing, as far as is possible, the need to be defensive about expressing deviant opinions and questionnaires usually include a rubric which aims to disarm the suspicious subject by assurances of anonymity or by statements designed either to disguise the true nature of the test or to imply

that there are ways of detecting 'false' answers as, for example, the Junior Eysenck Inventory does. But, in spite of this the objectives of several questionnaire items are discernible and may be prepotent.

Cattell recognises the influence exerted upon questionnaire authors by the need to construct questions so indirect that they gain some immunity to faking. Thorndike also sees the construction of items having low face validity as the most effective way of minimising deliberate distortion. However, such a procedure could produce its own disadvantages if respondents attempt to outguess the author and, because of the subtlety of his items, guess incorrectly. Furthermore, the problem of high scoring on socially desirable items may be a complex phenomenon incapable of satisfactory resolution since it may represent mere faking, a self-deceptive attitude, or it may indicate that a subject is genuinely above average on a socially desirable trait. The co-existence of these effects is a possibility and insistence upon external validation for each item and the disregarding of its apparent meaning is a possible route to a solution of this problem. This would not prevent idiosyncratic interpretations of items but would provide increased predictive validity through external referents.

2.45 Even when the honesty and conscientiousness of the respondent is not in doubt, the interpretation he places on the characteristics of each item can cause the accuracy of test results to fall. In order to report each response accurately the subject would need to isolate the single characteristic of a task or situation that the author is trying to depict when he uses a specific situation, e.g. 'watching team games', that possesses several characteristics.

In this context, a fundamental problem seems to be the provision, in pencil and paper tests, of adequate structure with respect to specific situational characteristics. As more and more details are incorporated into test items the resulting scores could increasingly become measures of reading comprehension.

Additionally, questionnaire scores can be influenced in many ways by the subject's linguistic and categorizing habits and his awareness of the attitudes and behaviours of others when he is asked for comparative ratings of his own actions as, for example, 'rare', 'occasional' or 'quite frequent' occurrences. Vernon attributes further inconsistencies among self-reports to the degree of suggestibility of a subject and concludes that highly educated persons such as students are often self-analytic and introspective, and more self-depreciatory than non-academics.

2.46 Another related concern is the extent to which the lack of explicit criteria against which respondents can scale their answers leads to decisions being made in borderline cases on the basis of response characteristics which are irrelevant to the aims of the test. For example, some subjects will tend more than others to avoid such alternative answers as "uncertain" or "in between", irrespective of the content of the items, just as some will acquiesce or tend to give a greater proportion of affirmative answers, while others adopt a positional response set, favouring alternatives which are located left, right, above, below, etc. Thorndike maintains that response set effects could be reduced by placing the respondent in a forced choice situation where "don't know" and "in between" responses are excluded. He also suggests that during the scoring of the test an alternative remedy might be

to tabulate such replies and employ a statistical correction. It is not clear what form of 'correction' could be applied to data containing a larger than expected number of intermediate responses. Indeed, such a pattern of scoring may indicate something meaningful about a respondent's personality: his cautious indecisiveness for example. The problems of positional and acquiescence sets seem to be less intractable and may be partly resolved by balancing the response modes so that as many affirmative as negative and 'left' as 'right' answers contribute to trait scores. Many of the personality questionnaires and inventories in current use adopt this method but Rorer (1965) suggests that there is little evidence that response style ever seriously affects questionnaire results. He makes a distinction between response style of which acquiescence is an example, and response set in which he includes social desirability. Cattell admits the existence of these effects and implies that questionnaires could be self-correcting, as has been attempted in form 'C' of the 16 P.F. Questionnaire. Referring to what he calls the 'new trait theory' he remarks:

> "This says that social desirability distortion and acquiescence response set, as well as other hypotheses about influences in faking and distortion, offer only a patchwork correction....the important principle is that raters, either rating themselves or others, bring about distortion which is predictable from their own personalities and roles." (1965, p.322)

He is here alluding to an earlier view (Cattell, 1961) that what earlier workers in personality measurement had regarded as flaws in their paper-and-pencil tests, were expressions of well-defined personality factors.

2.47 The specific ways in which distortion effects generally become apparent in different methods of personality assessment are

described as instrumental factors. For example, self-rated and other-rated estimates of what appears to be an identical trait in both tend not to correlate as highly as anticipated or, alternatively, correlation may arise between behaviours which do not belong to the same trait simply because they are measured by the same kind of method. Such a factor can be seen as a common element among variables which resemble one another in mode of presentation or of permitted response of scoring, but which does not extend to tests of the same construct that are couched in other modes. Cattell (1965) is optimistic about resolving instrumental factors which he described as single and therefore separable contaminating influences but Vernon considers this to be an overly-optimistic view and, while he agrees that instrumental factors will be revealed by properly designed experiments, he cannot accept that enough is known about their influences to enable their measurement and removal from personality data.

2.48 All these claims that questionnaires are sensitive to complex distorting influences lead one to ask whether they can be considered to be sufficiently authentic and stable for use in psychological research. On the other hand, it would seem to be foolish to reject such a convenient and easily scored instrument as long as it is recognised that it has certain weaknesses and accept it as one channel to an individual's conceptual system. When criticizing questionnaires it is well to keep in mind the availability of alternative means of personality assessment and their advantages and disadvantages. Such a comparison would yield several points in favour of questionnaires. Firstly, they can be easily standardised to yield norms which enable comparisons to be made between an

individual and others of his kind. Secondly, they usually incorporate a large number of items which have been shown by item analysis to relate to the central concept or trait and so tend to give a more reliable indication of this concept than, say, a few random questions at an interview. Thirdly, it might be that some respondents will be more frank and objective in answering an impersonal printed sheet than when writing biographical notes or being interviewed. Lastly, it is clear that any individual has access to infinitely more information about himself, his attitudes, opinions and feelings, than can possibly be obtained from other sources and, therefore, his self-ratings provide potentially greater accuracy than do alternative approaches.

2.49 An eclectic view about the vicissitudes of the questionnaire is held by Guilford (1959) who recognises that personality is so complex that we are unlikely to be able to subject all its aspects to objective measurement. He considers the psychometric approach of the questionnaire to be complementary to other methods: far from perfect but not meriting rejection if it can be shown to yield useful predictions.

2.50 The instrument used in this work is the 16 Personality Factor Questionnaire by Cattell & Eber (1967). It is widely used in clinical and research situations in Britain and the United States, and also in personnel selection by a growing number of university appointment services, colleges of further education, the public services and some of the largest commercial concerns (Jackson, 1974). In addition to the previously discussed problems which are associated with personality questionnaires in general there are some specific features of the 16 P.F. which deserve comment within this general context.

Each form of the 16 P.F. consists of 187 short statements or questions which require responses indicating agreement, strength of feelings or frequency of specified behaviours. Provision is made for responses to be entered in one of three ways for each item; the basic modes being 'yes', 'no', 'uncertain' with variations on these to allow expressions of intensity and direction of attitudes or relative incidence of behaviours.

Motivational distortion remains an obvious problem with some of the items. The authors' expressed intention was to develop the questionnaire to include more concept-valid and less face-valid items with the aim of reducing obvious belonging and vulnerability to faking. However, since the questionnaire is for application to the general adult population it is not possible to predict and prevent distortion occurring where items happen by chance to refer to a particular occupational group which includes the respondents, who may feel constrained to give answers showing orthodox opinions and expected behaviour. For example, the subjects in this research are teachers in training in a voluntary college of education. Some items refer directly to teachers, teaching, or have religious connotations and it is difficult to imagine that the subjects would not feel obliged to respond to these in an 'acceptable' way despite the rubric and assurances given by the tester that the results of the questionnaire could in no way be used to influence the student's progress in college. These are patent examples of probable sources of deliberate faking but many of the items contain subtle material which may also promote dissembling by subjects intent on playing the role of 'good student' or 'good teacher'.

2.51 The trend towards the inclusion in the 16 P.F. of concept-valid as distinct from face-valid items has been criticised as an

unsatisfactory way of resolving the problem of motivational distortion. The principal objection is to the authors' method of validating items against internal, factorial criteria only, which, according to Vernon (1963), does not eliminate social desirability effects but merely partitions them out among the various factors. This view reveals a more basic issue concerning not just the questionnaire itself but the model of personality upon which the 16 P.F. and kindred measures are built. As already outlined, Cattell's model relies upon the consistency of factor patterns obtained when data from varied personality measurement techniques are analysed by rotation to a criterion of oblique simple structure. This yields 'source traits' which are described as the major causal entities lying behind the more superficial clusters of associated personality variables. The criticism is that the validation procedure which correlates test items with factor constructs obtained in this way is inadequate.

> "The common element running through a set of inter correlated tests may be a response set, or halo or social desirability, rather than a presumed trait. In other words, factorial validation and internal consistency validation reduce to much the same thing as content or face validation." (Vernon, p.215)

2.52 On the other hand, Cattell argues that the conceptual validity of a test is its important property. Also the factor scales of the 16 P.F. do correlate meaningfully with concrete performances and if it were not so the factors represented must be written off as of no practical relevance. In support of this he lists some examples of source traits which correlate with

performance, viz:

Factor A	(affectothymia)	c. 0.5 with salary as a salesman
" B	(intelligence)	c. 0.6 with success in school
" C	(ego strength)	c.-0.4 with severity of neuroticism

It is interesting to note here that these examples use the factors which account for the maximum variance in personality measures.

These differences of opinion concerning validation do not appear to be irreconcilable but may be interpreted as different emphases on the routes to validation. Cattell considers the primary source of validity to be the correlation of items with factors consistently appearing in personality data, with confirmation coming from behavioural correlates, whereas Vernon stresses the importance of the items' correlations with external criteria and the patterns of scores in groups of people with well-known characteristics.

2.53 Both the strengths and weaknesses of the 16 P.F. arise from an attempt to gain the maximum amount of meaningful information about personality from a questionnaire measure, and herein lies the commonly-observed distinction between quality and quantity, which is compounded by the shortcomings of questionnaires in general. The 16 P.F. offers the possibility of obtaining multifactorial data to an extent not available in other tests, e.g. the Eysenck Personality Inventory, but in the process some sacrifice of reliability has been made. This is inevitable as even in a lengthy booklet the number of items which contribute to any single factor score must be small when a multiplicity of factors is being scored. Vernon takes the view that reliabilities of factor scores in the 16 P.F. are often inadequate: the test authors understandably take an opposing view and cite coefficients

for test/re-test on 146 subjects which range from .76 to .93 across the factors on two forms of the questionnaire.

One disturbing conclusion which Vernon comes to is that the 16 P.F. is too variable - it performs well in one investigation and not in another. The review of literature in which this test has been used tends to support this view but it needs to be qualified by taking into account that other variables providing data possibly less valid and reliable have been used in conjunction with the 16 P.F. There is little evidence to show that the source of unreliability is exclusively or even largely this instrument.

Personality and Teaching

2.54 Studies of personality characteristics and teaching competence have occupied researchers for some considerable time despite the enormous problems inherent in adequately describing the variables (Garner, 1973). Therefore it is not surprising that, as Getzels and Jackson (1963) report:

> "Despite the critical importance of the problem and a half century of prodigious research effort, very little is known for certain about the nature and measurement of teacher personality, or about the relation between teacher personality and teaching effectiveness. The regrettable fact is that many of the studies so far have not produced significant results. Many others have produced only pedestrian findings." (p.574)

Handley (1973) claims that two prerequisites of teaching success seem to stand out among the diversity of results generated by research in this area.

> "The first is intelligence, i.e. the power to discern and utilize relationships, without which good teaching would be impossible...the second...is a capacity for making human relationships." (p.86)

Both these attributes can properly be accommodated within most of the definitions and theories of personality reviewed earlier, and the influence of intelligence and other personality variables will now be considered. In very few of the reported studies does the researcher give details of the method of assessing the quality of teaching and, in the absence of this guidance, one is left to assume that either impressionistic methods or general rating schedules have been applied.

2.55 When, in 1931, Cattell published a list of twenty-two 'irreducible personality qualities' derived from his research, intelligence as a general factor, stood second in order of potency in accounting for psychological differences among individuals. However, Pinsent (1933) in an investigation into the antecedents of students in a university department of education, was unable to find any significant relationship between intelligence and teaching practice mark. Vernon (1939) studied the abilities of training college students and found no significant correlations between teaching ability and intelligence measured by both verbal and non-verbal tests. Researches by Lovell (1951) and Evans (1952) concurred. These findings tend to support the hypothesis that intelligence is not a determining factor in the success of students who show greater homogeneity with regard to this quality than does the general population from which they are drawn. The results obtained by Collins (1958) indicated no relationship between inadequate teaching ability and relatively low intelligence as measured by the AH5 Group Test, and Tarpey (1965) using the same test, was unable to find any link between intelligence and teaching

practice marks awarded by four colleges of education. The AH5 test was also applied by Wragg (1972) for comparison with Flanders' Interaction Analysis Data: only one correlation was significant and that was between the test and category 6 (gives instructions).

2.56 Results showing the opposite emphasis were obtained by LaDuke (1945) who used pupil gain as the criterion of teaching ability. Working with a sample of thirty-four teachers and two hundred students, he applied a wide range of instruments but concluded that only one predictor, intelligence, was significantly related to pupil gain. Rostker (1945) also found that intelligence was the most important factor in teaching ability but Rolfe (1945) in a replication of Rostker's experiment could not confirm his result. Carlile (1954) and Lomax (1969) used several measures of intelligence in their researches and obtained both positive and negative correlations with practical teaching, while Lomax's results seemed to indicate that intelligence operated differentially; women students of higher intelligence achieving better results on teaching practice.

The evidence for intelligence playing a major role in determining who shall succeed as teachers, once selection for training has occurred, is conflicting but much of it seems to suggest that it is of relatively little importance. The conviction that intelligence is important in teaching ability has been reinforced by common sense as has the notion that other personal qualities are influential. On such a basis researchers have been persuaded to undertake the study of the personality dimensions of student teachers and practitioners.

2.57 Barr (1948) summarized work concerned with the prediction and measurement of teaching efficiency. He reported more than two

hundred references to claimed relationships between personality and the criterion. Panton (1932) found significant correlations between some facets of personality and appearance of men students and their teaching marks. Some later investigators showed interest in the Bernreuter Inventory as a possibly useful instrument but Ward and Kirk (1942) reported very small correlations obtained from it in their study of the selection methods for entry to college and the eventual teaching performance of students. Conversely, Retan (1943) claimed that by using the Bernreuter in a battery of tests, it was possible to give as good a forecast of teaching ability in schools as could be obtained from teaching practice marks. One outcome of this work was the indication that unstable teachers functioned less satisfactorily after several years' service. However, some thirty per cent of the teachers who were rated as 'good' showed some signs of emotional instability.

2.58 Rostker (1945) could find no significant link between personality and competence but Rolfe's (1945) follow-up produced contrary findings. Von Haden (1946) examined ratings on eight personality traits among teachers and found significant correlations at the 1 per cent level for 'energy', 'initiative', 'professional judgement' and 'work habits'.

2.59 Using objective tests for each of Cattell's source traits, Schwartz (1950) concluded that none of the tests was related to the teaching success of the thirty-four students in his sample. Lamke (1951) used the 16 P.F. Questionnaire in conjunction with supervisory ratings of teacher effectiveness and concluded that those high school teachers who were rated as 'good' showed

higher scores on source traits F (surgency) and H (parmia). These teachers displayed qualities of being talkative, cheerful, placid, frank and quick, frivolous and emotionally responsive. They were also claimed to have strong artistic or sentimental interests and interests in the opposite sex. The poorly rated teachers were below average in these respects. Lamke concludes that the implications are that various (but certain) combinations of traits characterize good teachers; and various combinations of others, poor teachers. Additionally

> "Teachers are more or less successful not because they are cast in the same mould but because there is a kind of 'balance' among their personality traits." (p.247)

Montross (1954) also used the 16 P.F.Q. but obtained results which conflicted with those of Lamke. Only factor 'A' 'affectothymia' proved significant. High scorers described as warmhearted, easy-going, participating, were more often rated as successful teachers.

Warburton, Butcher and Forrest (1963) attempted to find predictors of the performance of a hundred students in a university department of education from data from the 16 P.F.Q. Their findings showed that higher scores on three factors: 'G' (superego strength), I (premsia) and Q_3 (self-sentiment) were significantly related to the criterion. The same questionnaire was used by Tarpey (1965) in a research to assess its predictive utility in student teacher selection. In only one of four colleges from which her sample was drawn did any of the source traits emerge as predictive. She found 'G' (superego strength), 'A' (affectothymia), 'H' (parmia) and 'M'- (praxernia) to be linked to teaching mark.

2.60 Chabassol (1968) found evidence of a basic difference in the relationship between successfully-rated male and female teachers and their pupils. Successful male teachers appeared to be more demanding, authoritarian and hostile, while successful female teachers tended to be warm and possessive. Chabassol draws the conclusion that inventories which hope to predict teaching success should take account of the sex of the teacher. Different patterns of qualities for successfully-rated men and women teachers also came from the study by Gough, Durflinger and Hill (1968), who used the California Psychological Inventory. Men rated as good teachers were conscientious, practical, rational, moderate and methodical. The qualities claimed for women teachers included dominance, perseverence, persistence and ambition.

McClain (1968) examined the 16 P.F.Q. scales which he thought might be related to teaching success in secondary schools. One objective of the study was the combination of the factor scores to assist in differentiating superior and inferior teachers. A mixed sample of almost two hundred students, just over half of whom were women, were rated on teaching competence. Correlations between test results and ratings were computed and weighted to maximize the predictive effect of factors taken in combination. Control, sobriety, steadiness, responsibility, non-competitiveness and freedom from tension were the hallmarks of successful men; and competence, energy, enthusiasm and spontaneity characterized successful women.

Henjum (1969) administered the 16 P.F.Q. to students who had been assessed on teaching practice in many different junior high and senior high schools. His criteria of teaching ability

were the usual supervisor's rating and the Hoyt-Grim Pupil Reaction Inventory. Significant correlations were claimed between teaching success and many personality factor scores. Different structures seemed to be shown by successful teachers among the two types of school. In junior high schools, significance was claimed for higher scores on factors A, B, C, D, E, F, G, H, I, M, N, Q_1, Q_2 and Q_3. These account for a great part of the second order factors which indicate extraversion and social adjustment. In the senior high schools success seemed to be dependent upon intelligence and enthusiasm.

2.61 Lomax (1969) used multiple regression analysis to evaluate data from a large battery of tests which he applied to sixty-eight men and women students stated to be representative of the academic departments and training groups in a college of education. Low scoring on factor 'M' (praxernia) was the only significant outcome from the 16 P.F.Q., indicating that more successful students were practical and concerned with facts.

Davis and Satterly (1969) applied the 16 P.F.Q. to 149 women students at their time of entry to training college and again, before the final teaching practice two years later. The personality profiles of small groups of students of high and low rated teaching ability were compared. Four factors emerged which differentiated the groups on both occasions of testing. These were: 'G' (superego strength), 'I' (tough-mindedness), 'O' (untroubled adequacy) and 'Q_4' (relaxed behaviour). Additionally, 'M' (praxernia) differentiated between the groups on the first occasion. The investigators suggested that poor performance was found when tender-mindedness, high insecurity, and tenseness were associated with a lack of conscientiousness.

2.62 Among the researches reviewed here there is a recurrent emphasis on certain personality factors' association with competent teaching: higher scores on 'G', 'M', 'A' and 'H' seem to be most frequently mentioned. The detailed popular descriptions of higher scorers on these factors are given by Cattell and Eber (1962) as follows:

G (superego strength)	Conscientious, persevering, staid rule-bound. Tends to be exacting in character, dominated by a sense of duty, responsible, planful.
M (autia)	Imaginative, wrapped up in inner urgencies, careless of practical matters, bohemian, tends to be unconventional, self-motivated, imaginatively created.
A (affectothymia)	Warmhearted, easy-going, participating, tends to be good-natured, emotionally expressive, ready to co-operate, attentive to people, soft-hearted, kindly adaptable. Likes dealing with people and socially-impressive situations. Readily forms active groups. Generous in personal relations, less afraid of criticism and better able to remember names of people.
H (parmia)	Venturesome, socially-bold, uninhibited, spontaneous, ready to try new things, abundant in emotional response. "Thick-skinnedness" enables him to face wear and tear in dealing with people and gruelling emotional situations without fatigue. However, he can be careless of detail, ignore danger signals and consume much time talking.

These four 'contributing' factors and the apparent contradictions among them are discussed in the summary.

A few studies have correlated personality variables with interaction analysis data in order to minimize the evaluative aspect of classroom observation. Austad (1972) reports only chance correlations between personality and observations in

microteaching bu Wragg (1972) found relationships significant beyond the 1 per cent level.

16 P.F.Q. factor 'A' and		F.I.A.C. 3	r = .25
		Teacher question ratio	r = .31
	'I'	F.I.A.C. 8	r = .27
		Teacher talk	r = -.37
	'0'	F.I.A.C. 6	r = .27

No very strong factors emerged from an analysis of a selection of the presage, process and product variables but one main factor - 'warm accepting' - loaded on the 16 P.F.Q. factor 'A', the accepting of feelings and ideas, questioning, pupil talk, and flexibility (i.e. the variety of lesson pattern).

In a study designed to relate classroom process criteria and teacher variables, Birkin (1971) obtained results indicating that among the best predictors, and those appearing most frequently in multiple regression analyses were factor Q_4 (stable - tense) and the second-order factor Anxiety - Adjustment. This last factor related negatively to the percentage of teacher-talk, the more anxious teachers talking more; and Q_4 also indicating that it was the tense teachers who operated a more direct approach, that is, had lower indirect-direct ratios among the interaction analysis categories.

Elements of Creativity

2.63 Within the last twenty years considerable interest has been shown in forms of thinking which are claimed to contrast with those involved in analytical problem-solving. As a consequence, much research has been done on modes of thinking and abilities involved in open-ended test situations where no single solution provides the criterion of unsuccessful performance. The gross

ability to adopt unconventional and productive avenues of thought when required has been labelled 'creativity'.

It seems appropriate to consider the concept of creativity and its more commonly hypothesised constituent elements, their relationship to that form of cognitive activity traditionally associated with the term 'intelligence', tests of creativity, and creativity and teaching.

2.64 Attempts to answer such questions as "What is creativity?" and "Who is a creative person?" have probably been made for centuries and have ranged within and between those areas now designated philosophical, psychological, sociological and statistical, and from the completely specific to the totally general. The word creative and its derivatives have been common currency as general and laudatory referents to qualities possessed by persons who have risen to eminence in various fields of endeavour. When used in such contexts they were rarely defined. Galton, in his study of 'hereditary genius' was able to apply the term creative without definition, with no reason to suppose that readers of his work would misinterpret its meaning; and Spearman (1929) declared that psychology could provide no adequate account of creativeness, again with little likelihood that his meaning would not be grasped. It should also be noted here, for later comment, that Spearman was convinced that it was one manifestation of general intelligence.

2.65 The development of interest in creative abilities has been described by Burt (1949), who highlighted the main themes and turning points and showed their relationship to the then current thinking. Among the important stages he mentioned is the work

of Galton and his students in the testing of 'higher mental processes'. Early in this century, Galton had abstracted, as components of the creative process, what he called 'special aptitudes' of receptivity, intuition or insight, and fluency ('an unusual and spontaneous flow of images and ideas'). This list of aptitudes was augmented by McDougall's proposal of 'productive or deviant association', and Garnett's (1919) 'c' or cleverness factor, which was concerned with quickness of apprehension, verbal humour and a capacity to produce unique responses. Hargreaves (1927) reported fluency of imagination to be a complex factor incorporating general intelligence, memory, speed and 'x'. The nature of 'x' and of speed being unclear but probably best described in conative terms as an absence of self-criticism or inhibition which might lead to a preference for quantity over quality in response situations.

Taylor (1947), in follow-up studies on the work of Thurstone (1938) found evidence for two distinct forms of fluency in expression: verbal and ideational, and concluded that since the ideational fluency involves the amount a person can communicate about a given topic and measures indirectly the flow of ideas, it is more fundamental than word fluency, which involves the handling of words solely in terms of their structure.

Hadamard (1945), in an essay on the psychology of invention, discussed the characteristics of creative people and concluded that artists often have a desire to do something different just because it is different. Houston and Mednick (1963) tested a similar hypothesis - that highly original individuals have a strong preference for novel responses as such, or an urge to avoid the

trite and the banal. Their experimental results tended to support the hypothesis, but it was difficult to state which version of it was the more tenable. This theme of novelty or originality as a component of creativity is a recurring one and is emphasised by Guilford (1950) who is credited with some of the earliest factorial studies,

> "The creative person has <u>novel</u> ideas. The degree of novelty of which a person is capable, or which he habitually exhibits, is pertinent to our study."
>
> (p. 452)

And again:

> "This can be tested in terms of the frequency of uncommon, yet acceptable, responses to items. This tendency to give remote verbal associations in a word-association test; and to give connotative synonyms for words, are examples of indications of novelty of ideas in the category of verbal tests."
>
> (p. 452)

This aspect of newness of thought is brought out in Drever's (1952) definition:

> "Creative: Producing an essentially new product, constructive (somewhat wider); used of imagination, where a new combination of ideas or images is constructed (strictly when it is self-initiated, rather than imitated; also of thought synthesis, where the mental product is not a mere summation."
>
> (p. 56)

This emphasis upon the spontaneous nature of the synthesis is also evident in Koestler's (1964) work based upon studies of many individuals judged to be creative in the 'understood' and 'accepted' sense. He sees creativity as involving 'bisociation': the application of unusual frames of reference to the solution of a problem. The routine skills of thinking are envisaged to lie in a single plane and 'bisociation' is described as always

operating in more than one plane

> "The former may be called single-minded, the latter a double-minded, transitory state of unstable equilibrium where the balance of both emotion and thought is disturbed." (p. 36)

2.66 One aspect of Koestler's view of creativity seems to be that of a problem-solving activity freed from constraints to adopt a single-minded approach. This flexibility is seen by Jackson and Messick (1965) in perceptual terms, to be a type of intellectual fluidity which is reflected in an ability to perceive objects in their own right - independent of their symbolic representation, their stereotyped function, or their relatedness to the immediate needs of the viewer. Similarly, Vinacke (1952) considered most thinking to alternate between two poles which he labelled the realistic and the imaginative. The realistic involved fairly strict adherence to logical, scientific criteria and being tied in one's responses to the external situation; the field is dominated by reason and facts. Thinking at the imaginative pole allows inner currents to play with the data originally provided by perception. Imaginative activity involves the thinker in fairly free experimentation with the data which throws up hypotheses, suggestions, fantasies, images and comparisons, and often strives towards unclear and barely conceived goals. Wright et al (1970) concur with Koestler and Vinacke in viewing creative activity in the context of problem-solving or goal-oriented behaviour and they see it as a special case where the originality and value of the solution are stress. Maltzman (1960) has, in fact, proposed a definition of creativity as 'originality evaluated'.

2.67 The tenor of the views expressed so far is that creative activity yields novel or original products which have value, and that this presupposes a flexibility as opposed to a rigidity of thought processes. The question of the value, or, to use Guilford's term, the acceptability of the product, is one of the conceptual problems examined by Jackson and Messick in their consideration of creativity assessment. They see the labels 'correct' and 'good' as applying differentially to the terms 'intelligent' and 'creative'. Intelligent responses satisfy objective criteria and are correct; they operate within the constraints of reality and logic and can be considered true or false, right or wrong. In contrast, creative responses are good; they satisfy subjective criteria although they may not be restricted by the demands of reality and logic. They are responsive to a wide variety of judgmental standards. These standards are given a more definite orientation by Haefele (1962) and Mednick (1962) who stress the social worth of the innovation, and by Parker (1963) for whom the criterion of goodness is the value the new product has for the innovator. The former view would exclude products from classification as original on the basis of bizarreness or contextual irrelevance and the latter brings in a dimension of self-evaluation which would not provide a stable basis for quantifying originality.

2.68 The evidence for the usefulness of the conceptual dimensions of fluency, flexibility and originality in providing a descriptive framework for creativeness has been amassed piecemeal from a range of studies, some of which have had little theoretical common ground in information collecting, processing or interpretation. But the recurrence of these dimensions and their

widely differing origins may be in themselves significant as Guilford (1947) conjectured when remarking upon disappointments in attempts to establish common factors of these types. It was Guilford (1956) who later, in his model of the intellect, was able to include such factors within a comprehensive theoretical structure which is the most elaborate attempt yet made to describe individual differences. After considering all the known factors that he regarded as belonging in the intellectual category, including the abilities of fluency, flexibility and originality as well as sensitivity to problems, he proposed a system of those factors. This 'structure of intellect model' classifies the factors in a three-fold way, demonstrating three principles by which they can be organised, namely: in terms of contents, operations and products, so that each primary intellectual ability represents an intersection of a certain kind of operation, applied to a certain kind of material, yielding a certain kind of product.

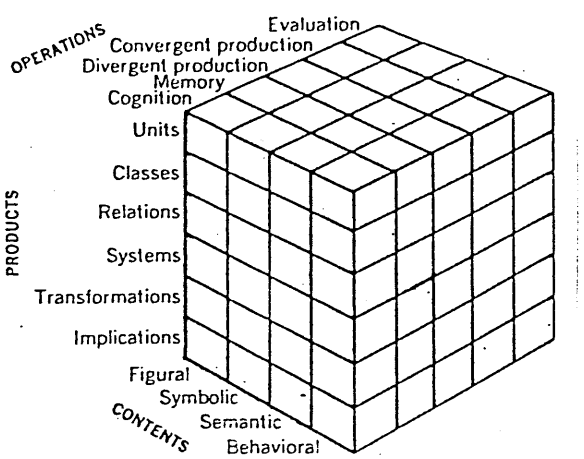


fig. 2

The traits of fluency, flexibility and originality come within the categories of divergent productions and transformations which

might serve as definitions of creative thinking although Guilford (1959b) is of the opinion that it would be incorrect to assume that they account for all the intellectual components of creativity.

Guilford's usage of 'divergent' and 'convergent' merits some brief explanation. He considers divergent thinking to emphasise searching activities with freedom to go in different directions whereas convergent thinking activities proceed towards one right answer, or one that is more or less clearly demanded by the given information.

2.69 Hudson (1966) uses the same terminology to designate individuals as convergers and divergers according to the dominant cognitive style but he denies that this distinction has any close connection with creativity. Hudson developed his work from the discovery that undergraduates at one university performed better on the non-verbal parts of intelligence tests if they were science students and more successfully on the verbal parts if they were reading arts subjects. A similar difference was found among able fifteen-year-olds and the bias was detectable in thirteen-year-olds before subject choices had been made in their schools. These points have relevance to the present investigation, since the emergence of contrasted cognitive styles will influence subject choices in schools and ultimately, as the self-selection is filtered through the examination procedure, the subject allegiance of students in higher education.

In debating the research maxim 'that creativity in all fields is associated with the same type - the diverger', Hudson assessed the finding of several studies of originality among adults by

Roe and by Mackinnon. The outcomes of these researches are taken to show that the relationship between divergence and creativeness is bound to be complex. In fact, the findings conflict. Roe reported that eminent research workers in physical science strongly resembled the converger: MacKinnon found creative men and women in all fields to be more divergent than their non-creative colleagues.

2.70 The main point at issue in these considerations appears to be the one of criterion which was stated at the outset: "What is creativity?" and "Who is a creative person?". Hudson is, at best, sceptical about psychometry and prefers to consider creativity in terms of a diversity of socially-recognised achievement; hence his emphasis upon success in the professions; Roe's investigation of scientists and MacKinnon's work with architects and other professionals, coupled with his own selection of evidence from the biographies of 'some great men' - Turner, Kepler, Darwin, etc.

This approach differs in essence from the factor-analytic standpoints of Guilford and Torrance, especially the latter with his stress upon the creative process rather than the product. This difference may be seen as a facet of the competence/performance gap where the possession of an 'ability' or a 'capacity' does not necessarily result in the appropriate production. This will be dealt with in more detail when tests of creativity are discussed but it is noteworthy that the criterion problem has been described as perhaps the most crucial one in this field (Taylor, 1964).

> "The criterion problem concerns the evaluation of the degree of creativeness of a product or a performance; it is quite separate from the

> prediction problem, in which the creative potential of people is estimated - for example, by means of test scores - and in which predictions about future creative performances are made, based upon the 'creative potential' estimate for each person." (p.9)

In subsequent modifications of his theory Guilford (Guilford and Merrifield, 1960) extended the net of creativity to encompass the 'redefinition' abilities in the convergent production category and 'sensitivity to problems' which falls into the evaluation category. By this action their concept of creativity was broadened to include re-cognition - literally re-knowing something familiar in a new way, and the openness or readiness to perceive possibly unconventional relationships. The research of Frick (1959) had indicated that too much familiarity with an object can work against a broader, more abstract view and Arnold (1962) quotes an example where employees in a manufacturing industry experienced difficulty in listing the attributes of their company's product. It is reported that Albert Einstein made a practice of denying his understanding of the obvious, in order that he might obtain a new look at things. These reports are consistent with the Jackson and Messick interpretation of flexibility in a perceptual mode and also with the Gestalt notion of the reorganisation or redefinition of organised wholes (Wertheimer, 1945). It is upon the extended concept of creativity that test batteries have been constructed by Guilford, Torrance and their associates to measure the thinking abilities involved. Among the factors and tasks the batteries typically include the following are most relevant to the present work (from Guilford and Merrifield).

Factor	Tests and Descriptions
Word Fluency	Suffixes: write words ending with a specified suffix Prefixes: write words beginning with a specified prefix First and Last Letters: write words beginning and ending with a specified letter
Expressional Fluency	Expressional Fluency: write four-word sentences when the first letter of each word is given Simile Interpretations: complete sentence that states an analogous idea Word Arrangements: write sentences containing four specified words
Ideational Fluency	Topics: write as many ideas as possible on a given theme Theme: write as many ideas as possible on a given theme Thing Categories: list the names of things that are round or could be called round Ideational Fluency: write names of things fitting into broad categories
Semantic Spontaneous Flexibility	Brick Uses (flexibility): write a variety of uses for a brick Alternate Uses: list different peculiar uses for common objects
Associational Fluency	Controlled Associations: write as many synonyms as possible for each given word Simile Insertions: write adjectival completion for a simile Associations: produce a word that can be associated with two given words

Factor	Tests and Descriptions
Originality	Plot Titles (clever): write clever titles for story plots Symbol Production: produce symbols to represent activities and objects Consequences (remote): list remote consequences of certain changes

Perusal of these factor names and related tasks reveals the important distinction which Taylor had made between the various types of fluency. The more highly-structured nature of 'associational' and 'word' fluency tasks with increasing specificity of permitted responses contrasts with the open-endedness of 'ideational' fluency: a contrast which is reminiscience of the convergence-divergence axis.

2.71 It is from Guilford's theory and tasks similar to those above that Torrance developed his approach to creativity assessment, firstly by adapting for younger children some of Guilford's materials and later by devising sets of tasks for application throughout the pre-school - post-graduate spectrum. This activity produced the Minnesota Tests of Creative Thinking which also incorporated material derived from studies of the characteristics of famous inventors. The major difference between Torrance and Guilford is the insistence by Guilford that predictor measurers should represent single factors, whereas Torrance has devised more complex tests each of which could be scored on several factors. Another point of difference is the emphasis placed by Torrance (1962) upon the process of creative thinking rather than upon the products which he sees as having so far claimed the major consideration. It is in this context that he defines creative thinking as the process of

sensing gaps or disturbing, missing elements; forming ideas or hypotheses concerning them; testing these hypotheses; and communicating the results, possibly modifying and retesting the hypotheses.

2.72 With the relatively recent and rapid increase of interest in creativity and its elements has come what was perhaps an inevitable desire to discover what relationships exist between creativity and intelligence as variously conceived. For some time there has been a growing suspicion that conventional tests of intelligence do not do justice to those people who have the capacity for imaginative and original thinking. Cronbach (1960) criticises such established tests and school examinations generally in that they have often failed to detect individuals whose subsequent careers have given proof of remarkable talent. He cites the case of Jan Masaryk who, as a child in the United States, was briefly confined in an institute for the mentally deficient as a result of his performance on an intelligence test. It is an indicator of this climate of opinion that the 1965 survey of eleven thousand children (Kellmer-Pringle et al, 1967) included among its measures a creativity rating but omitted any test of general intelligence. This zeitgeist has provoked considerable research and publishing activity amounting to what Hudson (1966) described as a bandwagon. The tide of literature relating to empirical studies of creativity, its nature with respect to intellect and achievement was heralded by a piece of researc by Getzels and Jackson in 1962. These authors were concerned that the notion of giftedness was generally regarded as synonymous with 'high I.Q.' and that this had come about as a

consequence of early enquiries tending to be restricted to the classroom, with academic achievement as their main concern. Getzels and Jackson sought to make comparisons between two groups of adolescent boys and girls in a Chicago private school: one representing individuals very high in measures of intelligence but not as high in measures of creativity; the other containing individuals very high in measures of creativity but not as high in measures of intelligence. The two groups were selected from 449 pupils on the result of either Stanford-Binet or Wechsler intelligence scales and five creativity measures taken from Guilford, Cattell or specially constructed for the study. Some verbal and ideational tests were included and consisted of word associations, uses for things, constructing endings for fables and making up mathematical problems capable of solution from given data. The scores on these and other sub-tests were summed to yield a gross creativity score.

The high-creativity group of the 26 students were in the top 20 per cent on the creativity tests but below the top 20 per cent in I.Q. The 28 students in the high-intelligence group were in the top 20 per cent in I.Q. but below the top 20 per cent on the creativity tests. The outcome of this research which has stimulated the most discussion and criticism relates to the interpretation of the finding that both groups were similar on academic achievement and, in particular, the implication that a lower I.Q. can be compensated by a higher creativity score. The further inference that creativity is only tenuously related to I.Q. and merits the status of a separate cognitive dimension was also made.

This work has been criticised on several grounds. Firstly, the high-creativity group, although described as having a low I.Q. had a mean quotient of 127 which places them almost two standard deviations above the population mean - a gifted group in the conventional sense. A second set of objections also bears upon the sampling which excluded from the study almost 400 of the 449 students, many of whom had performed well on both tests. The effect of this procedure was to create artifically separate groupings. Thirdly, and as Burt (1962) Wallach and Kogan (1965) point out, the intercorrelations of I.Q. and creativity tests were quite low but correlations among the various creativity tests themselves were not much higher.

2.73 Torrance (1963) replicated Getzels' and Jackson's study with pupils in eight junior schools and obtained results consistent with their findings. The overlap between Otis I.Q. and creativity amounted to only 30 per cent among those scoring in the top fifth of the sample on these measures.

Again, it should be noted that the sampling is open to the same criticisms as the previous authors'. A highly selective group, with a narrow ability range was used which tends to affect correlations between tests in a predictable way. McNemar (1964) on this point states that one need not be surprised at the fact that so-called creativity tests do not yield high correlations with I.Q., but the correlations are generally far higher than those found in typical studies with range restrictions. In these terms, the acceptance of creativity and I.Q. as substantially independent qualities from data produced by highly selective samples may be a statistical

artefact (Yamamoto, 1965).

Hasan and Butcher (1966) attempted a partial replication of the Getzels and Jackson study with a mixed group of children in their second year at a Scottish comprehensive school. The mean I.Q. of the sample was 102 and the test battery included four of the creativity scales used by Getzels and Jackson. Correlations among the creativity tests averaged 0.25 and between the intelligence and creativity tests 0.46. The results led the researchers to conclude that it might be possible to distinguish, by means of open-ended tests of divergent thinking, a set of children whose abilities have hitherto been concealed or minimised by the use of conventional tests but that this will not easily be possible in an unselected group of children by means of the tests when available.

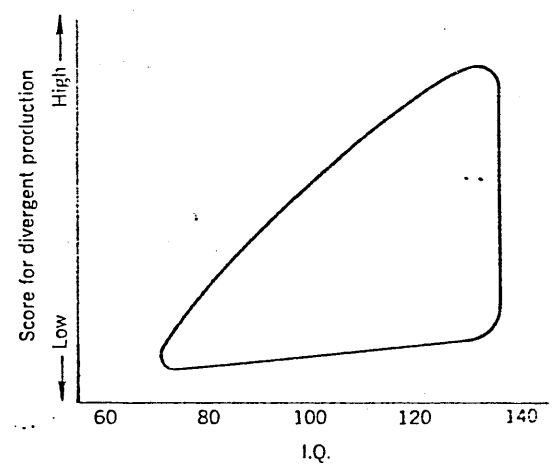
2.74 Another aspect of the controversy is that if the distinction between I.Q. and creativity is accepted, does whatever relationship that exists between them hold stable over a wide range of measures. This issue links McNemar's comment with the ambivalent results obtained, for if the relationship of creativity with I.Q. varies, the composition of the sample will be crucial. The few studies mentioned here and those reviewed in detail elsewhere (Vernon, 1970) indicate support for the Getzels and Jackson finding only where the sample is limited to subjects with very high I.Q. and little support where unselected groups have been used. Getzels and Jackson did seek to explain the intercorrelation of I.Q. and creativity scores by saying that this

"...signifies rather that a certain amount of

intelligence is required for creativity but that intelligence and creativity are by no means the same."
(p.215)

This threshold effect is underlined by McNemar who explained that at high I.Q. levels there will be a very wide range of creativity, whereas at average I.Q. and at lower levels the scatter for creativity will be less and less. Having a high I.Q. will not guarantee creativity and having a low I.Q. will make creativity impossible. The same conclusion is drawn by Guilford (1967) who relates the capacity for divergent production to I.Q. in graphical form.

fig. 3.



More high scores for divergent production are associated with high I.Q. than with low I.Q., but there are many moderately high scores on divergent thinking among those with average I.Q., and a high I.Q. is no guarantee of a high score on divergent thinking.

2.75 In summary, there are recurrent themes of fluency, flexibility and originality among reports ranging from the anecdotal to factor analytical describing creative individuals. A criterion problem remains in distinguishing between the concept of creativity

as a capacity or process and creativity as socially recognized achievement, that is, via created products. Although the problem of the relationship between creativity and intelligence as measured by tests has not been resolved in a generalizable statement, there is stronger evidence for the utility of creativity as a separate dimension when I.Q. is higher.

Creativity Tests

2.76 The testing of creative abilities is closely linked with the upsurge of interest in a broadening of the conceptual base of the intellect which followed Guilford's 1950 address to American psychologists. Torrance and his associates have developed Guilford's early work on testing and have given it a distinctive style which reflects the Minnesota school's view of creative thinking as a composite of abilities rather than a general factor of divergence. Although Guilford and Torrance are widely regarded as the authors and developers of creativity testing some related but more tentative investigations had been carried out by Taylor in the 1940's and by Thurstone even earlier.

Creativity tests vary widely in form and content. Some are based on verbal stimuli and responses, others are non-verbal in stimuli and responses, and some are hybrids of these modes. It is appropriate here to consider verbal tests since these form a part of the present work.

2.77 Guilford (1959) illustrates his concepts of divergent thinking abilities in a matrix whose rows and columns correspond to the products and contents axes of the structure of intellectual model. The ideational emphasis of the verbal abilities is recognized in

the labelling of the appropriate content column as semantic:

Kind of thing produced	Semantic content
Units	Ideational fluency
Classes	Semantic spontaneous flexibility
Relations	Associational fluency
Systems	
Transformations	Originality
Implications	Elaboration

Table 3.

(Adapted from Guilford (1959) p.382)

Ideational fluency, first remarked upon by Taylor (1947) is further defined by Wilson et al (1954) and Guilford and Christiansen (1956) as the ability to call up many ideas in a situation relatively free from restrictions, where quality of response is unimportant. Typical items in Guilford-type tests designed to assess this quality include namkin objects which meet a given specification, for example, things round and/or edible, and uses for common objects such as a paper clip. Scoring is independent on the number of answers produced in a given time.

Semantic spontaneous flexibility is defined as the ability or disposition to produce a diversity of ideas when free to do so. (Wilson (1954), Guilford, Frick et al (1956)). Test items intended to sample this take the form of questions about the uses of common objects, for example, a brick, where the flexibility score is counted as the number of different classes of use suggested. (Guilford considers this quality to be psychologically opposite to perseveration).

Associational fluency is illustrated by the individual's ability to produce words from a restricted area of meaning, for example, the writing of synonyms or completion of similes. Reports of this kind of fluency were made by Taylor (1947) and Fruchter (1948) although the concept was not as refined as in Guilford's use of the term.

2.78 Hargreaves (1927) is credited with the first report of originality in verbal data and this ability was given factor status in the Guilford model by Wilson (1954). Briefly stated, it is the ability or disposition to produce uncommon, remotely-associated, or clever responses. Suitable items for tests of this factor include the writing of unusual plot titles for a set of short stores and the listing of remote consequences of some hypothetical or unlikely event.

2.79 Torrance (1962) gives details of the test content of the Minnesota Tests of Creative Thinking which were precursors of the ones in present-day use and which bear his name. An inspection of the types of task the subject is asked to perform in these tests reveals the close similarity which exists between them and the earlier Guilford instruments.

Verbal tasks:

1. Impossibilities	List as many impossible things as you can.
2. Consequences	What would happen if ....
3. Just suppose	(Test content is similar to 2, but subject is given a drawing of an unlikely situation, e.g. just suppose clouds had strings attached to them which hang down to earth)
4. Situations	Subjects have to think of as many solutions as they can to such problems as (for children) 'If all schools were abolished, what would you do to try to become education?'

5. Common problems	Subjects are asked to think of as many problems as they can that might arise in ordinary situations like cooking a meal.
6. Improvements	Common objects, such as shoes, bicycles, etc. are given and the subjects are asked to suggest ways of improving them.

The above examples are not exhaustive of the range of items included in such tests but they indicate the development of the concepts of fluency, flexibility, originality, etc. in terms of their operational definitions. The approach to test construction by the Minnesota school differed from Guilford's in that tasks were constructed from an analysis of the experiences reported by eminent scientists, discoverers, inventors, writers, etc. The tasks themselves were intended to be models of the creative process by requiring the respondent to use several types of thinking rather than to produce a single factor score. Also, the authors tried to include a greater variety of stimuli; within the verbal forms of the Torrance Tests can be found verbal, pictorial and real object stimuli. The Product Improvement Test is an illustration of this approach. In this the subject is shown and allowed to handle a soft toy and asked to suggest ways in which it could be changed to 'make it more fun to play with'. Throughout the stages of development of the tests there has been a sustained attempt to extend their application; what began as an adaptation of Guilford's materials for younger children has yielded a set of tasks that can be used from nursery school to post-graduate level, in individual or group-test situations. The development of new tests was described by Yamamoto (1964) who refers to the tasks entitled: 'Ask', 'Guess Causes', and 'Guess Consequences'. In these the subject is required

to ask questions about a picture that cannot be answered merely by looking at it, and to make hypotheses about possible causes and consequences of the pictured behaviour. Again, these tasks are not entirely 'new', but the stimulus us and typifies the developing nature of creativity tests from this source.

2.80 As Foster (1971) remarks, creativity tests are as yet very unsophisticated measures of creative ability and this is indicated by the kind of criticism they have attracted. Comment upon the strengths and weaknesses of these tests seems to fall into four categories: content; administration; scoring and the univeral problems of reliability and validity.

Goldman (1964) criticizes some items in the Minnestoa Tests on the grounds of triviality, for example, questions about the uses of a tin can. One could take the view that it is unreasonable to expect subjects to be motivated towards creative production by tests involving such trivialities, if indeed they are trivial. On the other hand a widely-accepted attribute of creativeness is the ability to deal with the familiar in unfamiliar ways. The criticism may also be invalid if the age of the respondents is taken into account. Goldman admits that children and adolescents observed taking the test worked enthusiastically on items which adults may consider superficial. This raises the question of test-taking attributes generally which is beyond the scope of this treatment but the problem of getting the 'right' atmosphere for the test situation is a special one for creativity tests as we shall shortly examine. A further criticism of content arises from the derivation of items from Guilford's model. The remarks made

earlier about tests based upon factor analysis of the personality sphere apply equally well to these and need not be rehearsed.

The testing situation can be criticized on two main grounds: the 'atmosphere' may inhibit creative responses, and the possible effects of imposing time limits for the tasks. Torrance (1966a) instructs the 'examiners'

> "...to avoid the threatening situation frequently associated with testing. Create the expectation that examinees will enjoy the activities and invite them to 'have fun'. The psychological climate, both preceding and during the use of the tests, 'should be as comfortable and stimulating as possible'." (p.2)

Wallach and Kogan (1965) make the same point and develop it to include freedom from time limits and group administrations. The point at issue is a theoretical one articulated by Mednick (1962), which refers to the effects of time limits upon originality. The substance of this is as follows: if uniqueness is defined as a relative infrequency of a given associative response to the task presented to a sample of subjects, it is to be expected that stereotyped associates will come earlier and unique associates will come later in a sequence of responses. This view can be seen to pose problems for the tester if, as seems likely from a consideration of the concept of incubation, different time allocations may differentially affect scores on the various dimensions. The provision of a period for unconscious deliberation of the stimulus seems impossible to allow in test situations since problems introduced at one time and tested a few hours later may permit discussion among the subjects to take place or introduce other uncontrolled variables. A contrary view is

given by Taylor (1964) who remarks that in productive and creative thinking batteries of about sixty or seventy minutes duration, many individuals reduce both the quantity and quality of their production near the end of the time allowed. The most recent Torrance Verbal Tests have sub-tests of five or ten minutes duration amounting to forty-five minutes in all and it is the experience of the present writer that these time limits have been generous for the overwhelming majority of subjects. On this point Goldman (op.cit) concurs in stating that the time allowance is more than most children or adults will use.

2.81 The main problems associated with scoring the tests lie in the skills demanded by the complex and tedious methods which need to be applied in evaluating. All responses on verbal tests are written and each requires reading, interpreting and classifying in several ways. Of the three major dimensions, only fluency can be scored simply by a count of relevant responses. Despite this Yamamoto (1962) reports very high inter-scorer reliability for all three dimensions based upon 64 test records, the lowest correlation being .84.

Validity may be viewed in two ways: the effectiveness of predictions made from scores (predictive validity), and the internal evidence of factorial validity. Some predictive validation is provided by the work of Barron and Drevdahl (Vernon, 1962) although Anderson (1959) reports that predictive validity for creativity tests is generally poor. Anderson holds the view that creativity is often not suitable for predictive validity studies and prefers the construct validity procedures of which he describes exemplars. In the longitudinal study of

teachers by Torrance, Tan & Allman (1970) scores of originality were claimed to correlate satisfactory with life-style and teaching methods.

2.82 Harvey, Hoffmeister et al (1971) conducted a factor analytical evaluation of the Torrance Verbal Tests and found that some factor scores were more task-specific and showed less agreement with the hypothesised ability than the constructors claimed. This may be explained partly by the observation that not only are some persons creative because they are inspired by certain tasks, but also because one individual may be good at generating a large number of ideas (fluency) but poor in producing unusual (more original) ones. Another may be original within a limited area of application and yet may show poor flexibility in being unable or unwilling to step outside his frame of reference. Harvey and his colleagues, on the evidence produced by their study, suggested sweeping changes in the structure and scoring of the tests. They concluded that there is no justification for treating fluency as a factor distinct from flexibility and that two sub-tests; 'unusual uses' and 'unusual questions' should be eliminated as inadequate measures of fluency-flexibility and originality.

There is a practical dilemma here which exists independent of Harvey's work but is similar in kind. Torrance claims that what the tests measure are dimensions sufficiently different to make separate scores on them meaningful indicators of creative capacities and, although he does not recommend as a general practice the summation of these dissimilar scores to yield an

index of 'gross creativity',

> "Such a score does seem to give a rather stable index of the total amount of creative energy a person has available or is willing to use."
>
> (Torrance, 1966b, p.72)

Several of the researches reviewed in the previous section have shown no such scruples about the doubtful assumption of additivity of separate factor scores. At an obvious level the scores on fluency, flexibility and originality may be conceptually independent but they are mathematically related. The method of scoring the tests makes this inevitable since fluency scores have a limiting effect upon flexibility and, to a lesser extent, originality also.

Tests of creativity are in their infancy when we compare them with other psychometric instruments. They provide many pitfalls yet afford an opportunity to tap capacities not accessible to other tests. In the present study it is considered justifiable to include ideational measures to test hypotheses generated by the increasing volume of published work which suggests relationships may exist between fluency, flexibility, originality and teachers' classroom acts either evaluated or structurally observed.

Creativity and Teaching

2.83 At a common sense level the ability to be 'creative' or the possession of more specific qualities of fluency, flexibility and originality might be thought of as desirable in a teacher for none would feel praised if described by the adjectival opposites of these concepts: inflexible, unoriginal and having few ideas.

Investigations into possible relations between what are now called creativity dimensions and teacher performance have been less numerous than those concerned with intelligence and personality and it is interesting to trace their origins and development. Early work was concerned almost exclusively with verbal fluency and was stimulated by Thurstone's research into 'primary mental abilities' and the factor analytic methods applied in it.

2.84 Knoell (1953) and Montrose (1954) used similar tests of verbal fluency and found positive relationships between scores and teaching ability rated on a composite schedule. Their tests were adaptations of those devised by Taylor (1947) at the University of Chicago and the content of these included:

> Writing as much as possible on a theme
> Listing as many ideas as possible about a topic
> Sentence fluency - stating essentially the same thing in many ways
> Things round (number of round objects listed)
> Write as many adjectives as possible to describe a house
> Writing a clear, meaningful continuation of a story
> Listing as many things as possible to eat
> Completing similes.

Taylor claimed that in these tests, words are produced through a process of association, based upon meanings.

> "They are used as a means to an end, as tools in the expression of ideas. The ability is measured in some cases by the number of phrases or sentences produced. The words stand in some meaningful relation to other words, written or implied." (p.250)

Here we have an early creativity test in which the mode of apprehension and expression is exclusively verbal but the claim is that the instrument taps data which is essentially ideational. Knoell's hypothesis that ideational fluency would correlate highly with supervisors' ratings of teaching was substantiated and

it is possible to understand how supervisors may be impressed favourably by articulate student teachers who can present material verbally in novel ways as Crocker (1968) suggests. Also Beck (1967) shows that pupils see the ability to communicate as an important quality of a good teacher. The findings of Morgan and Woerdehoff (1969) link fluency scores with time-sampling counts of lecturing behaviour in Flanders' category 5, indicating a tendency to maintain a verbal exposition by the thirty-four student teachers they observed in secondary schools.

In the study of thirty-eight teachers by Knoell, the factor of ideational fluency (F) was significant at the .05 level and small positive correlations were obtained for two others: word fluency and verbal versatility. This last factor seems to be a compound of what other researchers have identified as verbal flexibility and originality.

2.85 A test of flexibility based upon items written by Guilford to assess divergent thinking was used by Crocker (1968) in a study of sixty-nine third-year college of education students, and this test emerged as a better predictor of tutor-assessed teaching success than were either intelligence or academic record. The ability to be flexible seems to be crucial if a student teacher is to cope with the problem that Ryan (1966) describes, that is, returning to a primary school and needing to assimilate new social mores and new role-relationships after having attended a secondary school for the last seven years or so. This comment places a somewhat different emphasis upon flexibility and it is possible to recognize the term's application

in two related areas of activity which may be conveniently labelled: 'pedagogic' and 'social', i.e. behaviour concerned with the means for organizing and communicating the material being taught, and the adaptability of the teacher to the pattern of inter-personal relationships demanded by the school. Indeed, Crocker (1973) comments that flexibility might merely be the ability to change teaching styles in order that the differing whims of class teachers and supervisors can be met. From an investigation of this link between teaching practice marks and flexibility scores among 391 students in two colleges of education Crocker concludes:

> "Provided we accept that the variety of words such as 'adaptability' are essentially being used to describe the same trait then this work would appear to support the researches of a great many other workers."
>
> (p. 425)

2.86 A few studies have reported connections between interaction analysis data and flexibility scores. The Morgan and Woerdehoff study found that flexibility contributed to multiple correlations between the test battery and F.I.A. categories 2 (praises or encourages) and 4 (asks questions), while Birkin (1971) discovered flexibility to be related to F.I.A. category 9 (pupil initiation). Both these researches applied derivatives of Guilford's and Torrance's tests. Again within the pedagogic frame, Joyce and Hodges (1966) identify flexibility with the capacity to show a wide variety of teaching styles and they believe that teachers who can are potentially able to accomplish more than those whose repertoires are relatively limited. Sprinthall et al (1966) recognize teacher competence as a dependent variable related to cognitive flexibility-rigidity as one of the most important

independent variables. Their study brings a fresh light to bear upon the discussion in that it has a psychodynamic orientation, the concept of flexibility being empirically based in the Rorschach and written versions of the Thematic Apperception Test. Their description of flexibility encompasses the pedagogic and the social emphases:

> "...the teachers' ability to think on his feet, to adapt teaching objectives, content and method in process (i.e. in response to the reaction, learning difficulties and needs of the pupils). More broadly... open-mindedness, adaptability, and resistance to premature conceptual closure...implies brightness, creativity and divergent thinking."

2.87 From a different orientation Torrance, Tan and Allman (1970) tried to predict the behaviour of 325 elementary school teachers by using scores on a test of creative thinking ability obtained during training, and self-reports on achievements and activities six years later. (The Tan Check-list). Complete data was collected for about one-third of the original sample. Teachers who had been identified as highly original students were stated to be living more fully and in their teaching they more frequently reported that they occasionally used role-play, problem-solving, panels, experiments, research, etc. On the other hand, those who had been identified as low originality students tended not to use these methods at all or used them regularly and continuously.

> "One also gathers the impression that variety and creative activities characterize their ('high originals') daily lives but that they are not compulsive about doing everything." (p.340)

This study also indicated that teachers who were 'high originals' were sensitive to others' responses to their original ideas. For example, they made fewer suggestions to supervisors and fellow teachers; the researchers' rationale for this being

that when original ideas are ridiculed, the person who is adept at producing them is less likely to make suggestions than are those low in originality. Such a sensitivity to problems is consistent with the earlier outline of the nature of creativity and the apparent compliance with the wishes and sentiments of authority figures can also be seen to be a source of approval for student teachers high in originality.

Summary and Research Hypotheses

3.1 The present work aims to examine the relationships between certain characteristics of student teachers and certain categories of actions performed by those individuals during teaching. The previous research reviewed seems to have produced a wealth of information that is not merely diverse but also fragmentary. Some reasons for this are made explicit elsewhere and need not be rehearsed but a prerequisite of any investigation in this area would seem to be the specification of the variables to be examined, together with a tentative theoretical framework of the instructional process within which to relate them.

A hypothesis derivation model has been proposed by Birkin (1971) who claims that it allows two important conditions to be met: first, a requirement that a wide variety of classroom conditions and variables to be taken into account in a clearly separable way; and second, that variables be differentiated in terms of such relations as dependent/independent and articulated one on another in systematically replicable ways.

3.2 A simplified version of Birkin's model is shown in fig. 4 to provide a frame of reference for locating the summary of points discussed earlier. The model is intended to suggest one possible approach to the analysis of teaching although it does seem to be sufficiently flexible to permit the operation of all the major directions of influence. For example, it is apparent that the model constitutes a closed system in that trends in the population variables, which reflect directive properties of past social

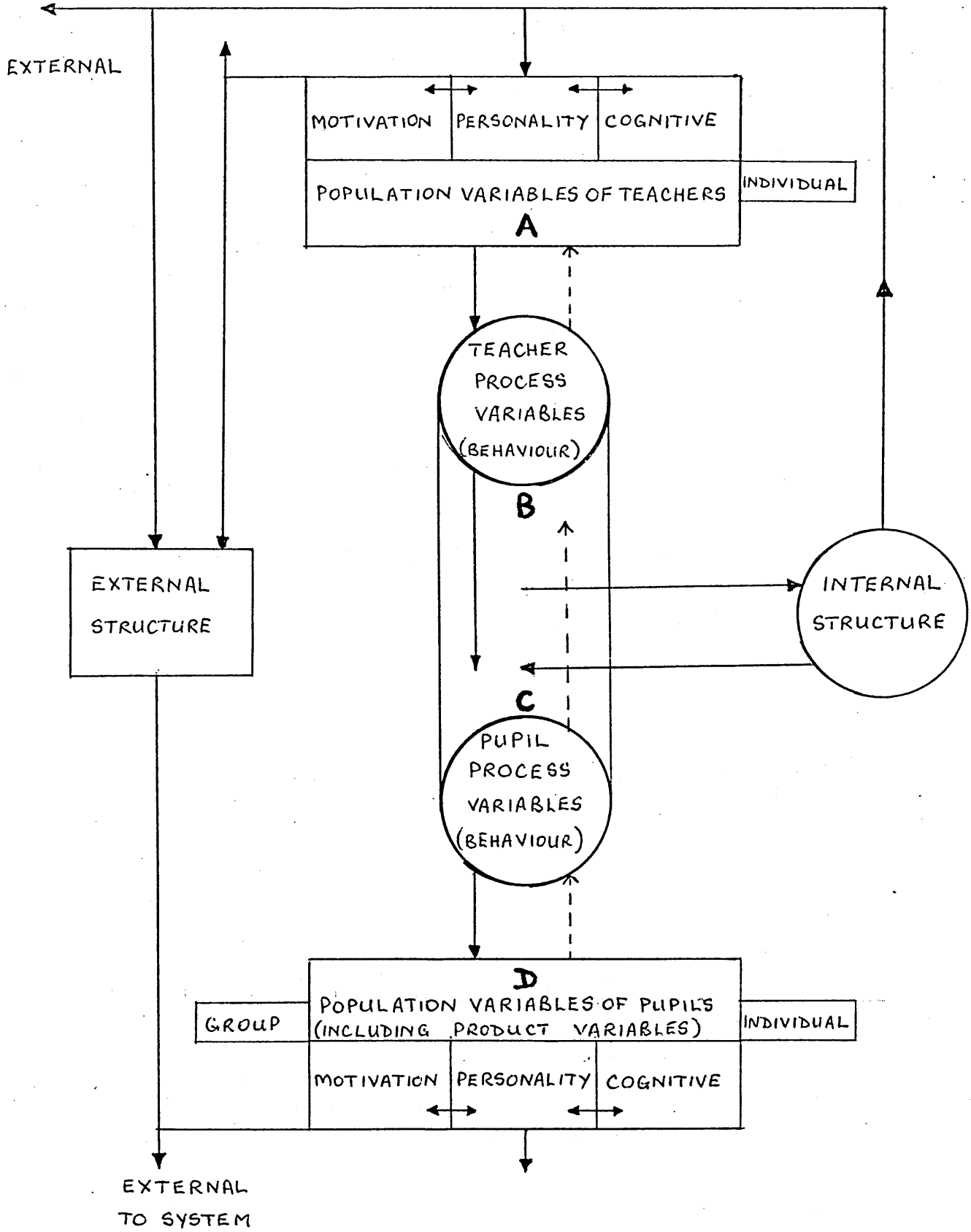

EXTERNAL
MOTIVATION
PERSONALITY
COGNITIVE
POPULATION VARIABLES OF TEACHERS
INDIVIDUAL
A
TEACHER PROCESS VARIABLES (BEHAVIOUR)
B
EXTERNAL STRUCTURE
INTERNAL STRUCTURE
C
PUPIL PROCESS VARIABLES (BEHAVIOUR)
D
GROUP
POPULATION VARIABLES OF PUPILS (INCLUDING PRODUCT VARIABLES)
INDIVIDUAL
MOTIVATION
PERSONALITY
COGNITIVE
EXTERNAL TO SYSTEM

fig. 4

experience outside the system, can re-enter it. Similarly, administrative factors which regulate the phasing of teacher and class contact, and physical elements such as spatial arrangements within the instructional setting are allowed their directive effects.

3.3 The model recognizes the predominant source of directed activity to be the teacher but does not exclude the reciprocal influences of pupils' actions. Its author acknowledges the unidirectional aspect revealed by the assumption that teacher acts generate pupil acts, and it is the consistent cycles of these that form the process of the instructional system.

3.4 If consideration of teacher success is limited to the appraisal of classroom acts the Birkin model shows up a major weakness common to many of the studies reviewed, that is, the inference of the nature and magnitude of changes brought about in the pupil population variables, which can be designated 'learning', from observations of teacher and pupil process variables. This would seem to be similarly suspect in all cases whether the observations are made impressionistically, rated, or recorded on a behavioural schedule, where there is an absence of demonstrably invariant relationships between points C and D.

When predictions of success are made from teacher presage variables such as attitudes, interests, personality or intelligence, the chain of assumptions is further extended when inferences about D are made from A.

3.5 The present work is concerned with a relatively restricted part of the model's topography: the relationship between some elements of A and B, and to a lesser extent C, since it is arguable that until reliable links can be established between

adjacent conceptual areas, more widely separated qualities and events cannot be satisfactorily related. However, this study cannot proceed without some assumptions being made in feeling one's way to the framing of testable hypotheses. The main problem rests in trying to link specific behaviours at B/C with teacher presage variables at A in a way that is consistent with the findings that certain personal traits correlate with teaching success. In other words, the investigator cannot avoid an evaluative component. He not only has to take note of the tendency for similar findings to occur in different researches but he has also to speculate why such behaviours have been adjudged 'good' or 'bad' practice and how these behaviours might bear a resultant relationship to the presage variable(s).

3.6 For the purposes of this investigation the description of personality given by Cattell will be used. A survey of the claimed relationships between primary source traits and teaching success shows the greatest consistency of findings to relate to factors A, G, H and M. In spite of the apparent contradictions among the descriptive labels of some of these factors - notably staid, rule-bound (G) vs. unconventional, bohemian (M) - it is possible to see a relationship between high scorers on these four factors and the claimed susceptibilities of tutors who assess teaching ability. Crocker (1968) concludes that the main influences at work are twofold: the relationship between the student and the pupils taught and the ingenuity shown in inventing new ways of presenting material to children. In terms of relationship with pupils, factors G, A and possibly H can be seen to be relevant in the personalities of teachers who would prepare

their lessons well, be demonstrative, talkative and slow to fatigue. The imaginativeness and spontaneity described by factors M and H would, if Crocker's opinion is valid, tend to impress tutors via the inventiveness and novelty of lesson preparation.

3.7 In this work it is hypothesised that higher scores on some personality factors will be significantly related to some categories of the Flanders' Interaction Analysis. The four factors just discussed will be considered first. It is felt reasonable to suppose that high scorers on factor A will contribute more often to categories 1 to 3 of the F.I.A. These are response categories in which the teacher is sensitive to pupils' verbal contributions and accepts and/or integrates them into the fabric of the lesson or gives praise or encouragement. This is consistent with the factor description of good-natured, attentive to people, kindly and adaptable. Flanders (1970) reports studies by Emmer in which the interdependence of interaction categories was investigated. Emmer found that the incidence of category 9 utterances (pupil initiation) was related to the amount of time devoted by the teacher to category 3 behaviour. It is a reasonable assumption, derived from operant conditioning theory, that the more a teacher reinforces pupils' contributions by building the lesson around them the greater will be the tendency for such contributions to occur. It is suggested that the category 9 data will also be positively related to the student teachers' scores on factor A.

The conscientiousness and dominance by a sense of duty of higher scorers on factor G may predispose them to careful lesson

preparation and execution. It is hypothesised that these qualities will also be revealed in inflexibility and a lack of spontaneity in the teaching and will result in more teacher initiatives, i.e. data in categories 4 to 6 will positively correlate with higher scores on factor G. The higher incidence of questioning by the teacher which is implicit in this hypothesis should evoke an increase in the pupils' responses in category 8.

Among the attributes of higher scorers on factor H are boldness, spontaneity and a tendency to consume time talking. It is possible that these traits would be manifested in an increased occurrence of lecturing behaviour and an elaboration of factual material and the teachers' own ideas. Within the F.I.A. data in category 5 are anticipated to correlate with higher scores on factor H.

It is difficult to suggest possible links between higher scores on factor M and the outcomes of F.I.A. since the quality of bohemianism, implying unconventionality and a carelessness of practical matters, does not seem to relate directly to any of the behavioural categories. On the other hand, low scorers on this factor tend to be careful, conventional and anxious to do the right thing. These descriptions suggest a similarity between the styles of behaviour of low scorers on M and high scorers on G, although the former are said to be regulated by external realities and the latter by moral constraints. It is hypothesised that low scores on M will positively correlate with verbal behaviours in F.I.A. Categories 4 to 6 and 8.

In view of the lack of consistently reported links between other personality factors and teaching performance no hypotheses relating to these are formulated here except to say that there is an expectation that factor B (intelligence) will not correlate significantly with any of the categories.

3.8 Of the ideational factors discussed in 2.63 to 2.87, fluency and flexibility have been claimed to correlate positively with teaching competence and with certain categories of F.I.A. Morgan and Woerdehoff's finding of significant relationships between lecturing (category 5) and fluency is consistent with the theoretical basis of this factor whether derived from verbal or ideational data. It is reasonable that the ability to generate a succession of ideas, or to give verbal expression to the same or similar ideas in different terms, should coincide with a discursive teaching style. For this reason it is hypothesised that this finding will be replicated.

The Morgan and Woerdehoff result relating flexibility scores with category 4 behaviour (asks questions) is consistent with Torrance's descriptions of this factor. It appears likely that a teacher who can generate a variety of hypotheses, shift from one approach to another, and use different strategies will also tend to ask more questions of the pupils. On these grounds it is anticipated that higher scores on flexibility will be accompanied by more category 4 and category 8 behaviours.

It is difficult to explain Morgan and Woerdehoff's finding of a link between flexibility and category 2 (gives praise) acts unless flexibility is conceived in an accepting rather than an expressive mode, yet the tests applied in that study do not

indicate this. A similar problem lies in Birkin's study where category 9 and flexibility were related. These authors do not attempt a clarification although in the former case the result is one of the most significant in statistical terms. No hypothesis is made here regarding the influence of flexibility upon category 2 and category 9 acts but it will be interesting to observe whether the results of the present work agree with these findings.

3.9 The research hypotheses to be tested are stated as follows:-

(1) Scores on factor A will correlate positively with categories 1, 2, 3 and 9 of the interaction analysis.

(2) Scores on factor B will not correlate with any category of the interaction analysis.

(3) Scores on factor G will correlate positively with categories 4, 5, 6 and 8 of the interaction analysis.

(4) Scores on factor H will correlate positively with category 5 of the interaction analysis.

(5) Scores on factor M will correlate negatively with categories 4, 5, 6 and 8 of the interaction analysis.

(6) Scores on Fluency will correlate positively with category 5 of the interaction analysis.

(7) Scores on Flexibility will correlate positively with categories 4 and 8 of the interaction analysis.

CHAPTER 4

DESIGN OF THE EXPERIMENT

4.1 The Sample

The subjects selected for study were students in a voluntary general college of education who were following courses for the Certificate of Education. All the students were in second-year training groups on a common Education course oriented towards the middle years of schooling, that is, towards the teaching of pupils within the nine to thirteen age range. The population was thus defined as those second-year students following the middle years course at this college.

4.2 In order to be able to make valid generalisations about population parameters it is necessary to consider principles of sampling, especially the principle of random selection under which each individual in the population has an equal chance of being chosen, and the selection of one individual is in no way tied to the selection of any other (Guilford, 1965). The criteria for selection in this study were twofold: (a) assignment, and (b) random allocation. Under (a) some students had been assigned on a geographical basis to the investigator for supervision of their teaching practice; each supervisor having a defined territory with a number of teaching places in schools within it. Students were assigned by a director of teaching practice solely upon their expressed preferences to teach pupils in a particular age group and, as this uniquely defines the population, it is thought justified to consider the assignment procedure as a lottery fulfilling the principle of randomisation.

4.3 Under (b) consideration of the informal hypothesis relating to sex differences among the criterion variables was instrumental

in ensuring sufficient numbers of men and women were included. This was achieved by noting the numbers of each sex assigned to the experiment under (a) and selecting a complement of students by drawing lots to arrive at approximately equal numbers over the two cycles of observation made in 1974 and 1975. These procedures are shown in fig. 5.

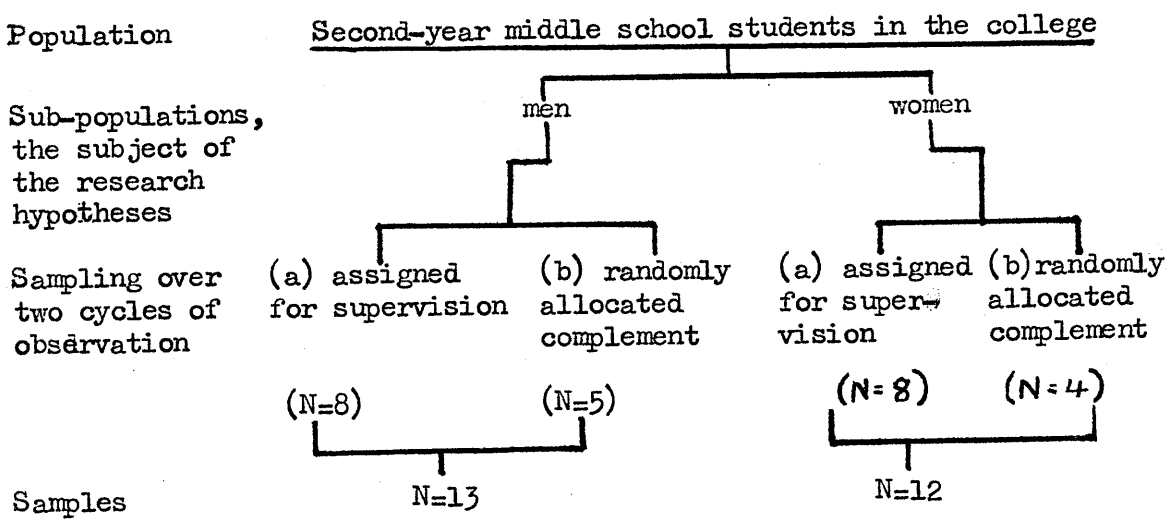


fig. 5

4.4 The choice of second-year students for study merits some explanation. It was made after considering the structure of the college course generally and the Education programme in particular. At the time of the investigation the teaching practices over the three year course were conceived as fulfilling different purposes. In year one the student's performance in a three-week practice was observed and reported upon with a view to providing appropriate guidance in teaching skills, so the purpose was explicitly diagnostic. However, the inexperience in the teaching role of a student faced with a complete school class for the first few occasions, coupled with the element of assessment and attendant anxiety, was thought to be a negative indicator for this kind of

investigation at that stage.

4.5 In year three an analogous situation arose where the student was well aware of the supervisor's role as assessor when grades for final teaching practice were being decided and might feel to be under greater constraint to please an observer. Only in year two was the teaching practice relatively free from formal assessment. During this year students' progress in their various courses was monitored mainly by continuous assessment and students were encouraged to treat the second teaching practice as a time for some experimentation and testing of their own curricular ideas, based upon their increasing confidence in the classroom. Their school experience at that time consisted of a half day per week group practice (one tutor with several students teaching groups of children within a class) for half a term, and a block practice of three weeks followed by a period of study practice (a combination of observation and teaching carried out during one half day per week over one term at the beginning of year two).

4.6 Additionally, the content of the Education course at the end of the second year posed a problem which might have been influential in the present study had third-year students been sampled. Brown (1975) reports that knowledge of an observational schedules such as Flanders' can materially affect a teacher's classroom behaviour without his participation in a training programme in which it is applied, for example, microteaching. That part of the Education course already referred to included a study of the Flanders instrument and its applications and for this reason also the sample was restricted to second-year students.

4.7 The sampling procedures selected groups of students analysed by their main subject of study as follows:

Table 4

First cycle 1974

	Main subject	Number of students
Men	Divinity	1
	Geography	1
	Physical Education	4
Women	Art & Craft	1
	Divinity	1
	Geography	1
	History	2
	Physical Education	1

Second cycle 1975

Men	American studies	1
	Art & Craft	2
	English	1
	French	1
	Geography	1
	History	1
Women	American studies	2
	Divinity	2
	English	1
	Music	1

The sampling fraction represented by these figures was approximately 1:7 in both cycles.

4.8 Lomax (1969) comments upon the representative nature of his sample in terms of the main subject affiliations of the students, and other researchers, e.g. Crocker (1973), have also considered these as variables potentially related to classroom behaviours. One can see possible reasons for the interest in subject affiliations of students: the criteria for choice of course or for acceptance by subject departments may be related to personal characteristics as suggested by Hudson (1966); methodology learned in a subject

context may influence a student's general teaching style, both of which would operate at point 'A' in Birkin's model. An inspection of table 4 reveals an over-representation of Physical Education students, especially men, and an under-representation of students of English, in terms of the composition of the middle years courses, while some subjects, e.g. sciences and drama do not appear at all. However such an imbalance is not an intentional or incidental feature, it has arisen by chance as the randomisation processes show. A degree of experimental control could have been introduced by stratified random samplings of sub-populations of main subject groups but this would have created problems of a logistical nature in that the investigator would have had to observe a sample which was largely distinct from the group assigned for supervision, and this would, of necessity, be an even smaller number of students.

4.9 The Schools

A serious limitation in classroom research is the problem of varied teaching situation. Questions about the validity of descriptions of classrooms and their comparability pose enormous difficulties for the researcher who aims to obtain generalizable principles from his work. This has led to a tendency among investigators to ignore or to gloss over differences. As Getzels and Jackson (1963) point out in their survey:

> "Investigators seems for the most part contentto take their subjects where they can find them, implying that a teacher is a teacher whether his school is in the country or in the city.... in an upper-class suburb or in a lower-class slum." (p.575)

These authors outline other gross sources of variation from one school to another such as educational viewpoints, the "nature"

of the pupils, conditions in the community related to teacher status, and they conclude that little provision is usually made for taking into account the effects of such situational variables on experimental results. There is, however, little firm evidence available on how these effects should be taken account of, in what directions and to what extent they are active.

4.10 Hall (1970) suggests some factors which affect a student's performance on teaching practice. Two of these are relevant here:

(i) the organization of the school, often with a group of newly-qualified teachers who have not yet solved all of their own teaching problems.

(ii) the pupils, with differences in social class and previous educational experience.

With the variables described by Getzels and Jackson these clearly fall into two areas of the Birkin model: pupil population and internal structure. What effects differences in these areas have upon teacher-pupil interaction remaining a matter for speculation but there is some evidence (Stones and Morris, 1972) that supervisors are prepared to make allowances for the "difficulty" of the school determined by the level of facilities, school environment, the number of 'problem' children, the degree to which the regular teacher has difficulty coping, and the co-operation and sympathy of the staff.

This last point is taken up by Crocker (1973) who tells how it is not uncommon for students on their first practice to be left entirely alone by the class teacher and yet to find on subsequent practices that a different teacher with different ideas spends almost the whole time in the classroom.

4.11 The 'school environment' is often a euphemism for the socio-economic status of the pupils in the school catchment area and this is often felt and claimed to relate to a student's development as a teacher. Ryan (1966) describes student teachers as sociological strangers in schools and his investigation showed a marked tendency for them to have unrealistically high expectations of their pupils. Crocker sees this problem aggravated where pupils are drawn from mainly low socio-economic areas since, according to Veldman and Peck (1963), pupils are more likely to rate students as poor teachers under these circumstances and this will affect the ways in which they respond to teaching.

4.12 This very brief introduction aims to touch upon some of the factors thought to be influential in promoting variability among schools which could weaken comparative studies of classroom behaviour. These variables are specifiable only in gross terms and at different levels of generality, for example, cultural, organizational and inter-personal, and as such cannot be controlled or 'allowed for' in the experimental design sense. It would be facile to claim that the schools used in this work were a random sample of those in the region for students to teach the middle age range of pupils in, but there are many kinds of constraints which operate to invalidate such a claim. The schools available were allocated to the college by the University of Liverpool A.T.O. on varied criteria, some geographical, some based on traditional and/or denominational links which antedate the university connexion. For example, available schools in well-defined areas of Merseyside, urban Lancashire and North Cheshire are predominantly Local Authority,

Aided or Controlled Church of England Schools, with a negligible number of schools of other denominations. Another constraint rests in the college policy of zoning teaching practice schools to achieve practical economies in the supervision of students and this applies to the assigned portion of the student samples discussed at 4.2.

There is therefore no foundation for claiming true randomization of schools in this study but so far as the population defined is concerned any incidental bias is a geographical one. It now remains to outline the characteristics of these schools.

4.13 The student samples were allocated to schools in the City of Liverpool and the adjacent towns of Wallasey, Huyton and Widnes in the Merseyside Conurbation. Seventeen schools were used and the composition of the set was as follows:

L.E.A. Junior Schools	7
C. of E. Aided Junior Schools	2
L.E.A. Middle Schools	3
L.E.A. Comprehensive Schools	5

Of the L.E.A. Junior Schools, five were late nineteenth century structures located in the older quarters of the towns and served areas of terraced housing dating from before the First World War and the inter-war period, while two of these were in clearance and development areas where a considerable amount of demolition of old property was in progress. Another two schools had been erected during the early and middle 1950s, mainly to serve the developing council housing estates, with a smaller proportion of owner-occupied houses.

The two church schools provided a great contrast; one was

built before 1870 and drew its pupils from almost exclusively, rather modest, private housing dating from various periods of growth ending in the early 1970s, while the other was a modern structure whose catchment comprised new local authority housing still undergoing construction.

The three middle schools also showed variety. Two of these were formerly secondary modern schools of the post-war period, one situated in a suburban council estate, the other near to the town centre with a wide range of types of accommodation to draw upon. The third had developed from a nineteenth century grammar school and drew pupils from mixed but predominantly owner-occupied property built prior to 1939.

Two of the comprehensive schools were amalgamations of modern and grammar schools on different sites located in mixed residential areas while a further two were modern buildings similarly placed. The remaining comprehensive school was a three site institution in a decaying quarter of the city adjacent to the docks. This had previously functioned as one co-educational and two single-sex modern schools.

4.14 It is not feasible to enumerate all the differences thought to be significant in any description of this group of schools but the schools themselves are apparently a fair cross-section of those available for teaching practice. Their internal organization could be described as traditional, that is, class teaching in the junior and specialist teaching in the secondary schools were the norms, with middle schools following a hybrid pattern.

4.15 The Personality Questionnaire

According to its author, Birkin's model of the instructional process imposes certain conditions on the design and statistical characteristics of research, in particular the application of multivariate procedures with the opportunities they offer for the manipulation of concepts represented operationally by patterns rather than by single variables. This imposition concerns both the measures applied and the methods of computation used to test hypotheses, and these will be described in turn.

The personality instrument most frequently used in recent British research is Cattell's 16 P.F. Questionnaire discussed in 2.50 to 2.53. This most nearly fulfils Birkin's requirement in providing scores on several personality dimensions which permit an examination of traits in combination and allow an exploration of possibly synergistic effects in the presage-process relationship. Anglicised forms A and B of the most recent (1970) editions of the questionnaire were applied before the start of each cycle of observation, one week intervening between the two applications.

4.16 In the writer's experience, students are sensitive to any suggestion of incidental assessment and it was felt necesssary to discuss wich each student, in somewhat guarded terms, the nature of the work being undertaken. The gist of the explanation given was that the work was being done in a private capacity and had no connexion whatever with the college's courses or assessment procedures. No information of any kind derived from questionnaires or observation of a student would be communicated in a way that would make possible the identification of an individual. In discussion each subject was t old that participation

was entirely voluntary and care was taken to avoid suggesting that a refusal would cause inconvenience to the investigator. Two students in the first cycle did express some unease and were replaced by random allocation.

4.17 The questionnaires were applied following the recommended procedure as closely as possible, the sample being tested in groups of six or seven at a time. From the scored questionnaires data on primary factors in raw and standard ten (sten) form and second-order factors were obtained, and these are listed in Appendix I. The authors urge that, wherever possible, two forms of the test should be applied, particularly in research and in all cases where maximum prediction is required. Also, forms 'A' and 'B' are described as the most suitable for college students. It is from this practice that Cattell & Eber (1967) derived reliability coefficients ranging from .76 to .93 across the factors on a sample of 146 over a six day period.

The trait stability coefficients are also relevant where a fairly long period of time elapses between applying the questionnaire and completion of the observation cycle. Some traits like F (surgency), M (imaginative), Q_3 (self-discipline), and Q_4 (drive tension) are more labile and can vary significantly over a few months. These coefficients for the factors across forms 'A' and 'B' ranged from .63 (factor B) to .88 (factor H) when computed from the test/re-test data gathered over a two month period on 132 subjects. In the present work testing and observation was completed within six weeks in each cycle and function fluctuation effects should therefore be minimal.

4.18 The Creativity Tests

Verbal forms 'A' and 'B' of the Torrance Tests of Creative

Thinking (Torrance, 1966a) were administered before each cycle and a few days after the personality questionnaire. Some difficulty was experienced in establishing the recommended relaxed game-like atmosphere among the participants, mainly it is thought, because of the time limits imposed which seemed to encourage competitiveness to produce a large number of responses, although the quality (cleverness and unusualness) as well as the quantity of answers was emphasized in the rubric. The time limits themselves were found to be generous; very few students were still writing their replies when time was up and most admitted that they had exhausted their ideas. A further property of the tests which militates against the required climate is the tedium associated with some of the items, especially the first three tasks in each form of the test where the subjects are invited to ask questions, guess causes and guess consequences of the action depicted in a cartoon. Some display of irritation was observable among the students performing these tasks, a frequent comment being "What? Again?"

In discussing the tests after the final administration several students remarked that they were "silly" and "could provide no useful information". These reactions are consistent with Goldman's (1964) criticism of triviality in the Minnesota Tests.

With these practical reservations in mind it is considered unsafe to accept uncritically Torrance's figures for test reliability and validity particularly in the light of the findings by Harvey et al (1971). It may be that the samples tested by Torrance and his associates did not regard the items as tedious and/or trivial and, in this case, one might raise questions about national differences in test-taking attitudes; Wragg (1972) has expressed reservations

about the suitability of the Torrance Tests for use in Britain.

4.19 Scoring the tests raised some problems of classification of responses especially in evaluating their originality, where the available scoring range was 0-1-2 and rather vague instructions are provided. Also, the flexibility categories are only apparently distinct and there is scope for some subjective interpretation in classifying responses. In the writer's opinion this instrument is the least satisfactory of the measures used in this present work but it is fair to admit that the author and publishers of the tests are aware that the assessment of creative thinking "cannot have reached the level of technical excellence that is eventually desired for it", and that the publication of the tests in their present condition is intended to

> "encourage research, facilitate data gathering and accomplish the very widening of knowledge in this area that is so urgently needed" (Torrance, 1966a, p.1).

The obtained scores for fluency, flexibility and originality are tabulated in Appendix II.

4.20 The Interaction Analysis

This part of the work demanded the acquisition of skills of observing, categorizing and recording events as they occurred in the classroom. Flanders (1970) prescribes a self-training programme of from four to twelve hours' duration to achieve proficiency in scoring reliably. This programme was worked through during a period of teaching practice preceding that used in the study and was continued until the observer could sort the tallies into categories with a minimum of hesitation and felt reasonably confident that the process of recording was not interfering with judgements of the interaction itself. An example of the type of tally sheet used is given in Appendix III.

In the gathering of data for the study at attempt was made to circumscribe the events judged to be valid for inclusion in the interaction analysis. For example, some lessons contain long periods of individual work where pupils may be reading, writing, or in some other way not interacting with the teacher; or the teacher may be reading a story, carrying out classroom administrative procedures such as marking the register, issuing or retrieving material. None of these situations represents the preferred interaction style of the teacher which would be reflected in a less-structured discussion, so to avoid scoring episodes like these the investigator asked each student to submit lesson plans to clearly indicate the nature and sequence of each activity. In this way it was possible to observe each student teaching in the same curriculum area, e.g. history, geography, topic work, etc. on each visit.

4.24 In order to allow a settling-in period for the students in the schools, no scoring of classroom events was done until one week after the teaching practice had started in junior schools. In the secondary and middle schools this was extended to two weeks to ensure that the student had taught the class at least twice before scoring commenced.

Flanders (1970) considers episodes of interaction which cumulatively yield four hundred tallies provide a minimum source from which meaningful statements can be made about a teacher's behaviour. In this work the aim was to promote reliability by exceeding this figure by as wide a margin as possible, given the constraints on conditions of observation, geographical distribution of students and the duration of the practice. In the event, episodes of interaction amounting to between 25 and 55 minutes and represented by approximately

500 to 1100 tallies, were achieved. The collection of the interaction data required the observation of each student on two or three occasions and details of conditions of these observations with analyses of the tallies in each behavioural category are given in Appendices IIIa and IIIb.

4.22 Treatment of the Data

The independent variables for each subject in the sample were:

(a) 16 P.F.Q. scores on primary and second-order factors (18)

(b) Scores on Torrance Tests of Creative Thinking (3)

The dependent variables for each subject in the sample were:

(c) Percentage tallies in each category of F.I.A. (10)

(d) Percentage 'Teacher Talk' (1)

(e) Teacher Response Ratio (1)

(f) Teacher Questioning Ratio (1)

(g) Pupil Initiation Ratio (1)

Variables (d) to (g) were not directly concerned with the research hypotheses but were included to reveal broad stylistic tendencies in teaching. Teacher talk is computed by summing the percentages of tallies in categories 1 to 7 and provides an index of teacher domination of the verbal interchange; the TRR is found by adding category frequencies 1 + 2 + 3, multiplying by 100 and dividing by the sum of 1 + 2 + 3 + 6 + 7, and this indicates the teacher's tendency to react to the ideas and feelings of the pupils. The TQR is described by Flanders as an indicator of the tendency of a teacher to use questions in guiding the more content oriented part of class dicsussion and is

dependent upon the extent to which the teacher solicits pupil responses to ideas which the teacher considers important or as he tests their understanding by asking questions. This ratio is computed by multiplying category 4 frequency by 100 and dividing by the sum of categories 4 + 5. The PIR shows the proportion of pupil talk interpreted by the observer to represent acts of initiation and is obtained by multiplying the frequency in category 9 by 100 and dividing by the sum of all pupil talk. Details of these indices derived from the interaction analysis data are tabulated in Appendix IIId.

4.23 The aspect of the experimental design which permits the summation of data gathered on separate occasions of observation assumes no significant changes in the proportionate distribution of time to the ten categories. This assumption was tested by analyses of variance of the data shown in Appendix IV.

Table 5.

Analysis of data gathered on two occasions:

Source of Variance	SS.	df.	MS.
between occasions	0.1	1	0.1
between categories	28077.4	9	3119.7
occasions x cats.	42.7	9	4.8
within cells (error)	6572.3	280	23.5
Totals	34692.5	299	

$$F = \frac{\text{MS. occasions x categories}}{\text{MS. error}} = \frac{4.8}{23.5} = 0.2 \text{ (n.s.)}$$

Table 6

Analysis of data gathered on three occasions:

Source of Variance	SS.	df.	MS.
between occasions	0.3	2	0.15
between categories	23374.1	9	2597.15
occasions x cats.	55.7	18	3.1
within cells (error)	12756.6	270	47.2
Totals	36186.7	299	

$$F = \frac{\text{MS. occasions x categories}}{\text{MS. error}} = \frac{3.1}{} = \frac{0.07}{} \text{ (n.s.)}$$

The results indicate no significant interaction between occasions and categories and hence no quantitative change in the classroom behaviours encoded.

4.24 Product-moment correlation coefficients between each independent and each dependent variable were computed for the two sub-samples (men and women) in order to test the research hypotheses. The intention here was to discover whether any single predictor held a direct relationship with a behavioural category beyond the conventional levels of statistical significance. The obtained matrices of correlations are given in Appendix V.

4.25 In order to consider possible relationships between the criteria data and sets of predictors a series of stepwise multiple regression analyses was performed, one for each F.I.A. category and for the derived ratios for each sub-sample, making a total of 28 computer runs. To effect these the Open University programme IDA (interactive data analysis) was used. The capacity of the matrix for this programme is limited to 19 columns of variables including the criterion, and as the present work includes 21 predictors, some editing of the data had to be made. The three variables having the lowest product-moment correlation with the criterion in each case were omitted from the analysis. This procedure is not entirely satisfactory because it is recognised that variables having very small coefficients can play an influential part as suppressors in multiple correlation (Guilford, 1965). The magnitude of the omitted coefficients ranged from .00 to .19 with a mean of .05 and an inspection of the regression analysis print-out showed that where coefficients of this order had been selected by the programme for inclusion

their contribution towards multiple R was generally quite small.

4.26 Each regression analysis proceeds stepwise by first selecting the independent variable having the highest correlation with the criterion and then by choosing successive variables in order of descending 't' value; that is, the measure of the null hypothesis that the corresponding b coefficient should be zero and that the variable should therefore not be included in the regression. The 't' value for each selected variable is recalculated at each step up to the limit of multiple R equals unity. Summaries of the computer print-out are given in Appendix VI to show the variables in the order selected on each run together with that proportion of criterion variance attributable to each calculated from the equation provided by Guilford (1956) p.397.

$$R^2_{1.23\ldots} = \beta_{12.3\ldots} r_{12} + \beta_{13.2\ldots} r_{13} + \ldots$$

Where R^2 is the coefficient of multiple determination,

$\beta_{12.3\ldots}$ $\beta_{13.2\ldots}$ etc. are the coefficients in standard form for each variable,

r is the product-moment correlation coefficient of each variable with the criterion.

CHAPTER 5

STATEMENT OF RESULTS

5.1 The principal statistical technique used in this work is the computation of product-moment correlations and the most important requirement for the legitimate use of the Pearson r is that the trend of relationship between the correlated variables be rectilinear (Guilford, 1965). Any marked departure from a straight-line regression may be detected by inspection of a scatter diagram and, in stating the results which follow, a plot of points for each pair of variables subject to specific research hypotheses has been made. In no case is any pronounced curvilinearity observable in the array.

5.2 Hypothesis 1.

fig. 6. sub-sample of men

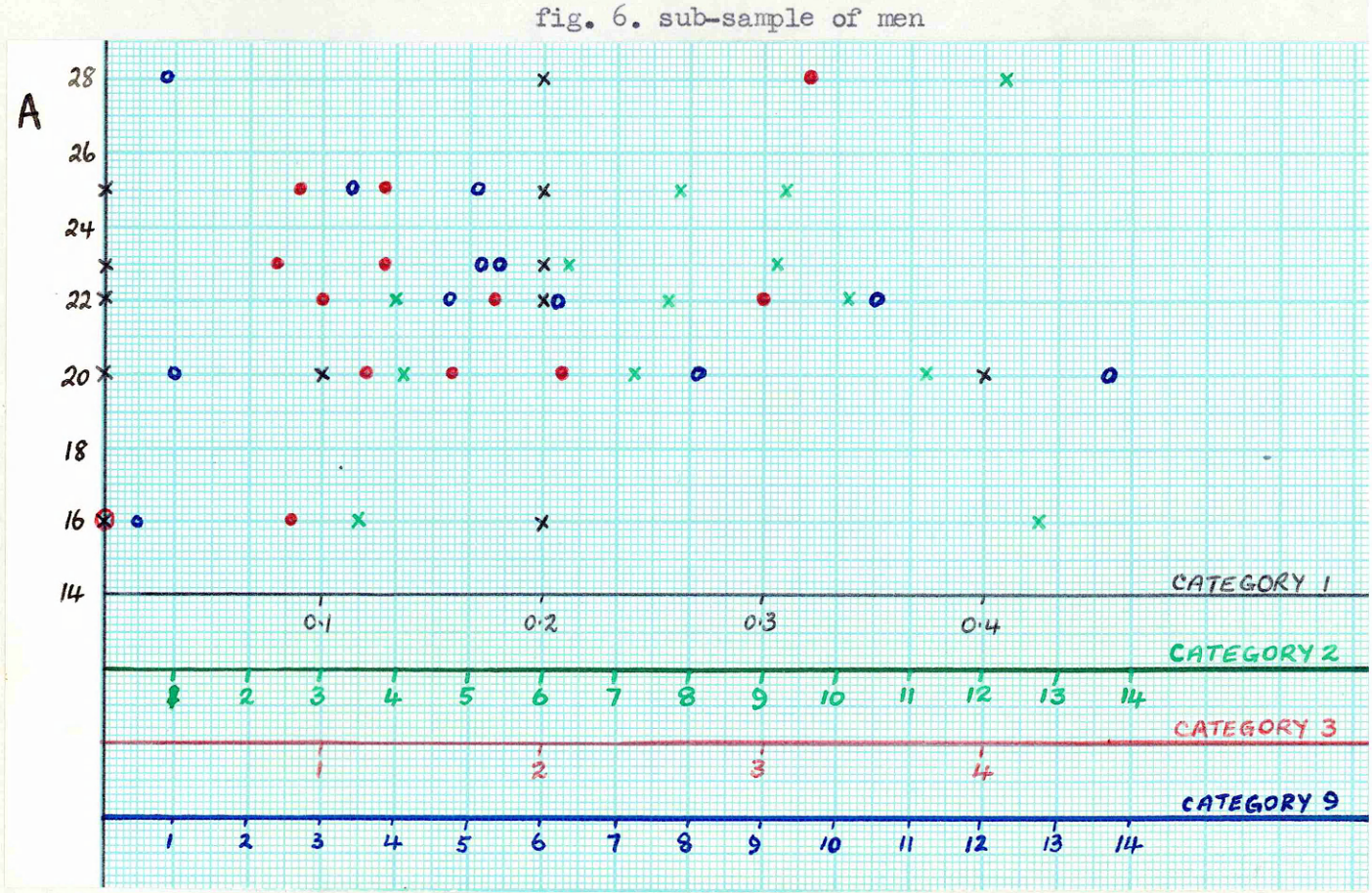


Table 7. product-moment correlations

F.I.A.C. %	1	2	3	9
Personality factor 'A'	-.13 n.s.	.41 n.s.	.48 n.s.	-.15 n.s.

fig. 7. sub-sample of women

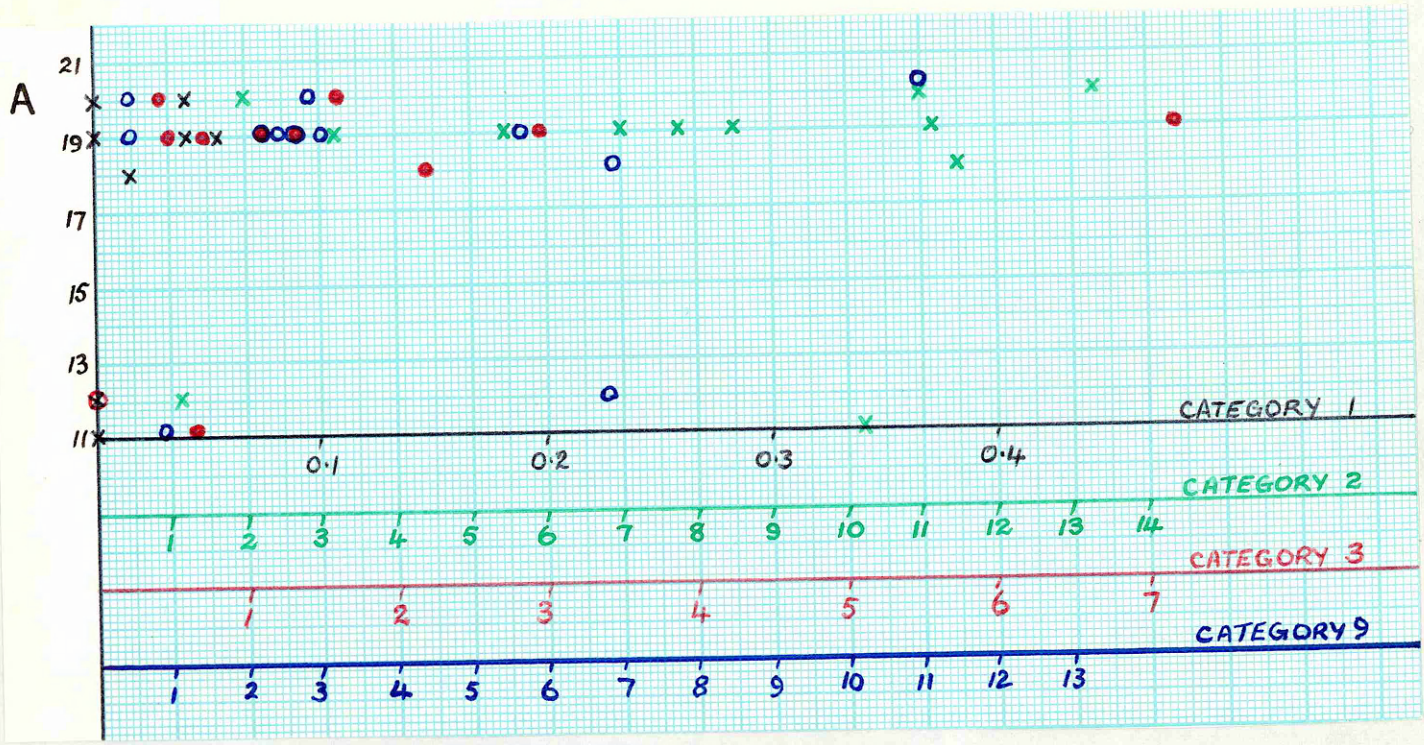


Table 8. product-moment correlations

F.I.A.C. %	1	2	3	9
Personality factor 'A'	.27 n.s.	.19 n.s.	.36 n.s.	.01 n.s.

Hypothesis 1 is rejected

Table 8A product-moment correlations (men)

F.I.A.C. %	1	2	3	4	5	6	7	8	9	10
Personality factor 'B'	03 n.s.	17 n.s.	07 n.s.	-22 n.s.	16 n.s.	25 n.s.	12 n.s.	-20 n.s.	-07 n.s.	-51 n.s.

Table 8B product-moment correlations (women)

F.I.A.C. %	1	2	3	4	5	6	7	8	9	10
Personality factor 'B'	16 n.s.	-12 ns.	-42 n.s.	02 n.s.	38 n.s.	-42 n.s.	-21 n.s.	-10 n.s.	14 n.s.	-35 n.s.

5.3 Hypothesis 2

In no case does the correlation coefficient for the relationship between personality factor 'B' scores and percentage occurrence of F.I.A.C. behaviour reach the magnitude associated with the conventional levels of statistical significance. (p.05; p.01) See Tables 8A and 8B

Hypothesis 2 is accepted.

5.4 Hypothesis 3

fig. 8. sub-sample of men

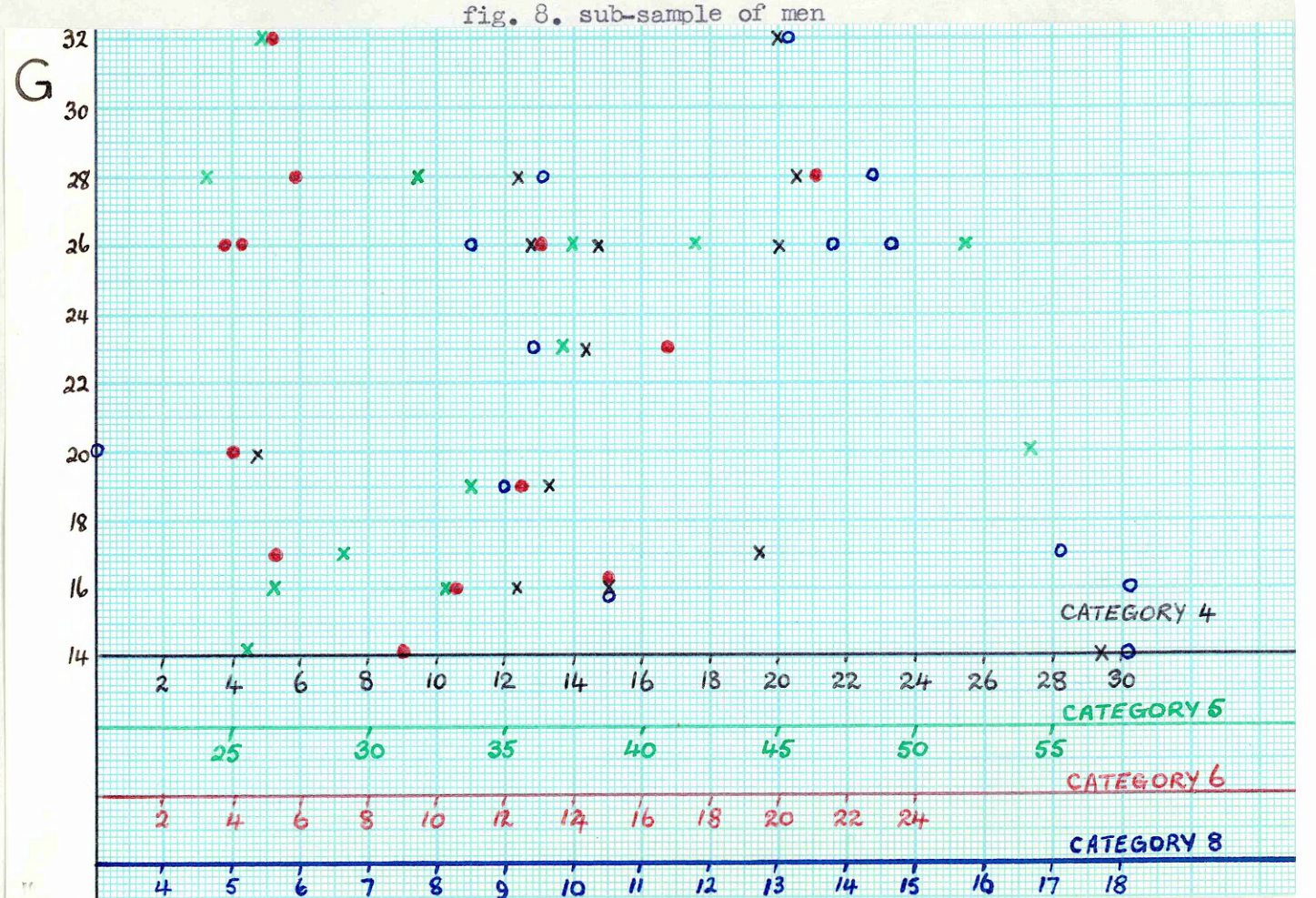


Table 9. product-moment correlations

F.I.A.C. %	4	5	6	8
Personality factor 'G'	-.07 n.s.	.10 n.s.	-.09 n.s.	-.22 n.s.

Hypothesis 3 - contd.

fig. 9. sub-sample of women

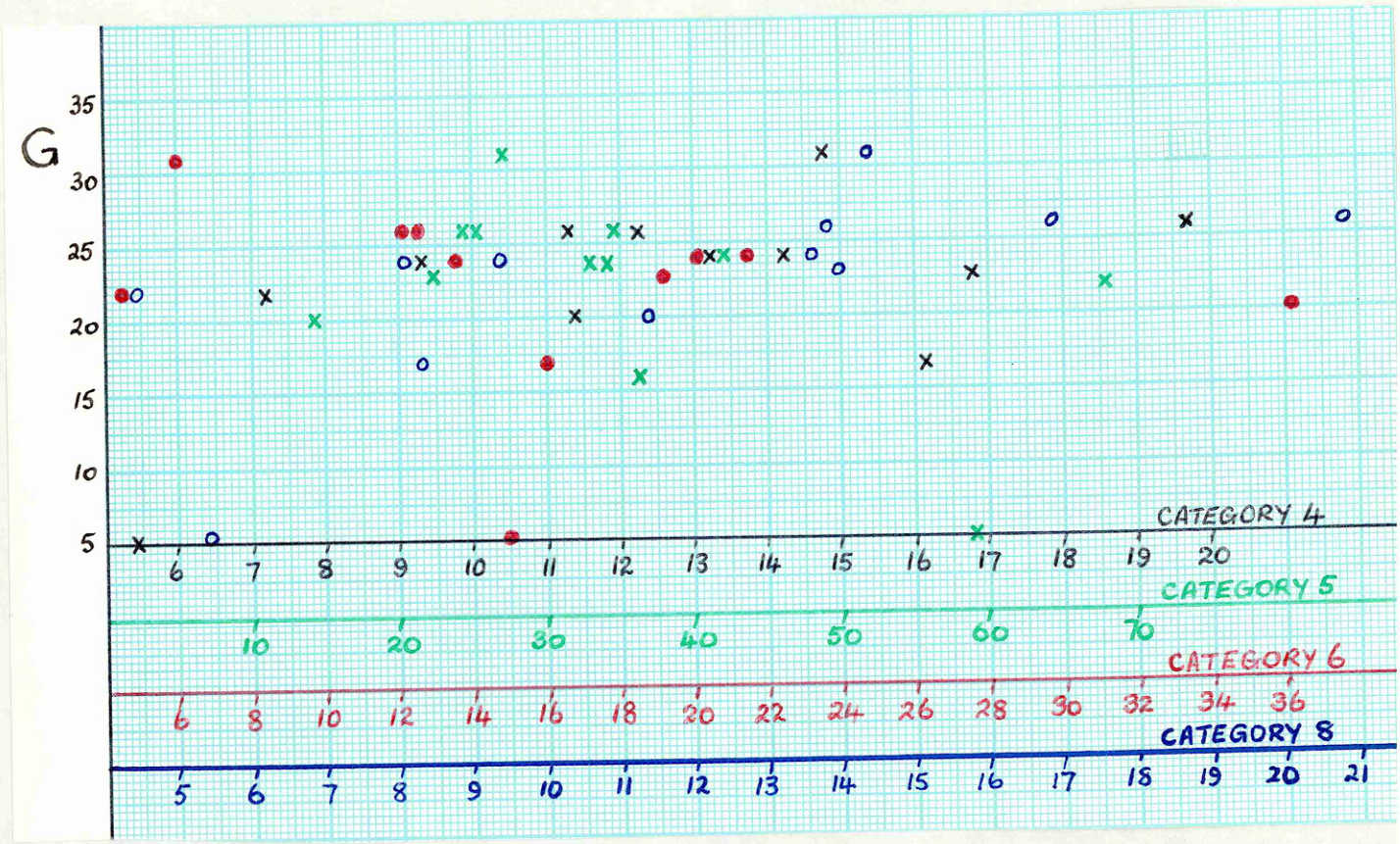


Table 10. product-moment correlations

F.I.A.C. %	4	5	6	8
Personality factor 'G'	.51 n.s.	-.49 n.s.	-.28 n.s.	.61* ~~n.s.~~

*significant at .05 level

Hypothesis 3 is accepted insofar as it applies to category 8 behaviour in women, but otherwise is rejected.

5.5. Hypothesis 4

fig. 10 sub-sample of men

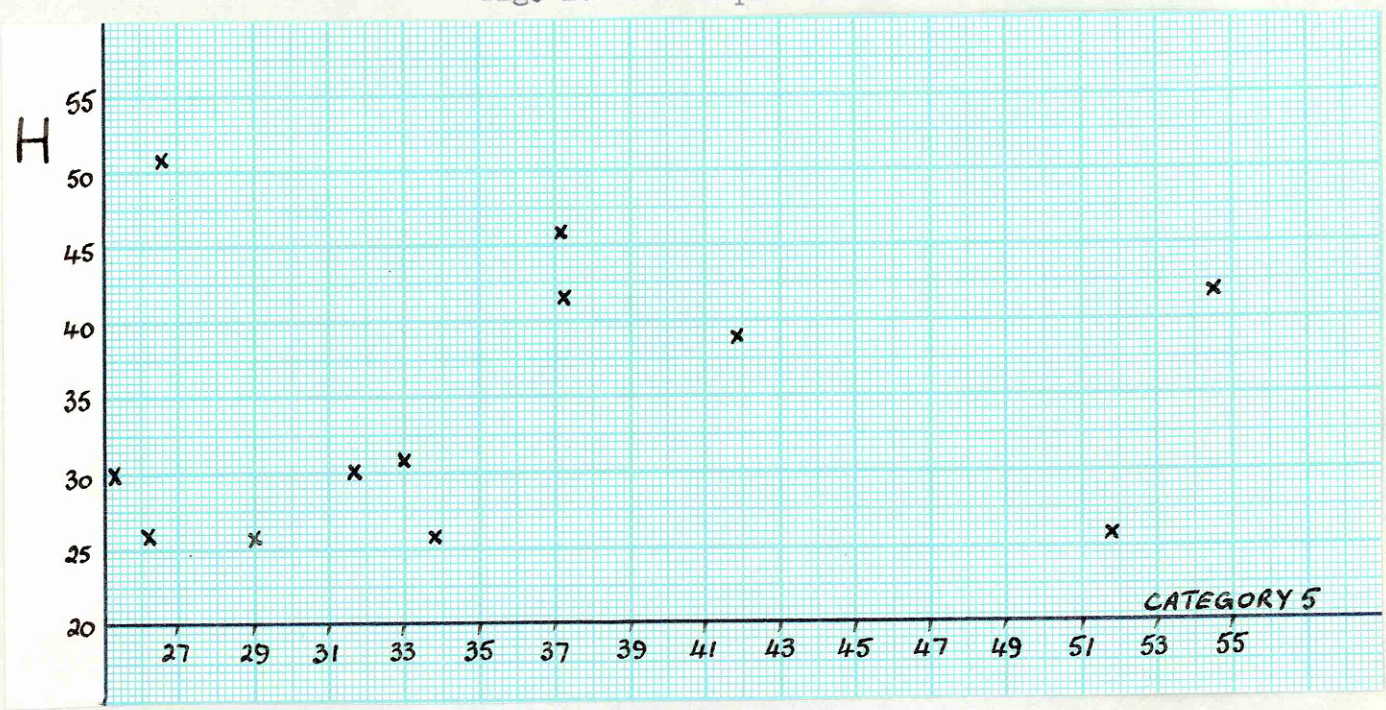


fig. 11. sub-sample of women

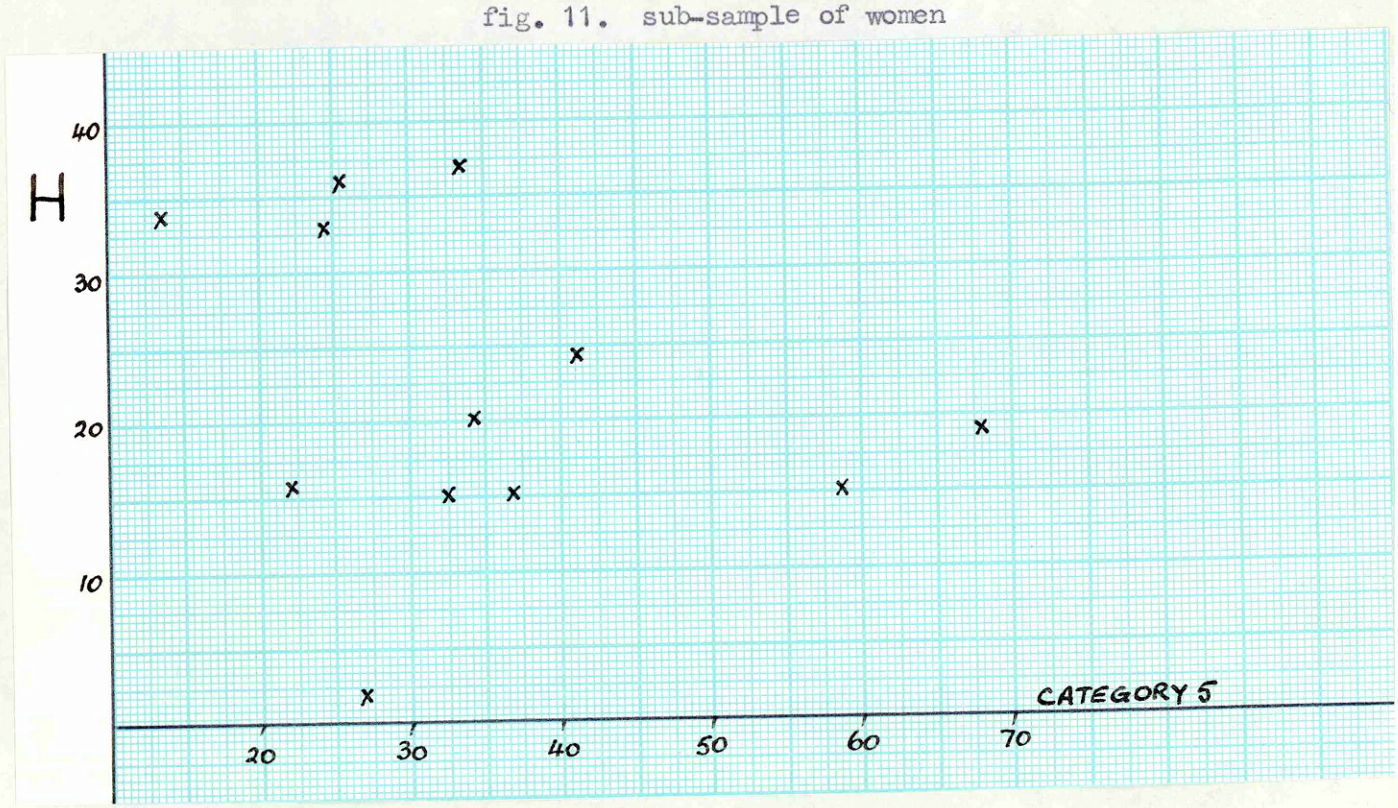


Table 11. product-moment correlations

F.I.A.C. %	5	
Personality factor 'H'	.25 n.s.	(men)
	-.30 n.s.	(women)

Hypothesis 4 is rejected

5.6 Hypothesis 5

fig. 12. sub-sample of men

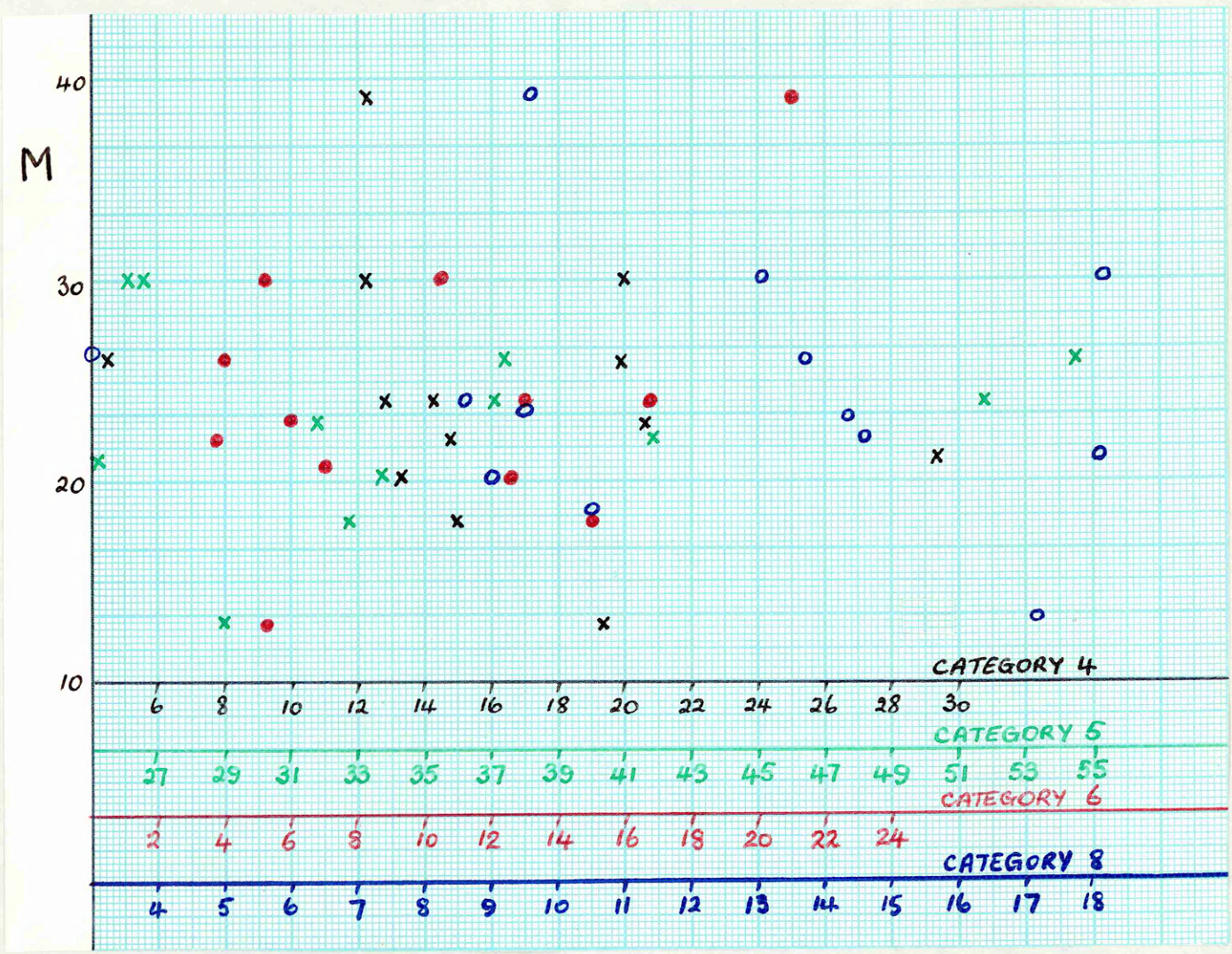


Table 12. product-moment correlations

F.I.A.C. %	4	5	6	8
Personality factor 'M'	-.29 n.s.	-.17 n.s.	.34 n.s.	-.21 n.s.

Hypothesis 5 - contd.

fig. 13. sub-sample of women

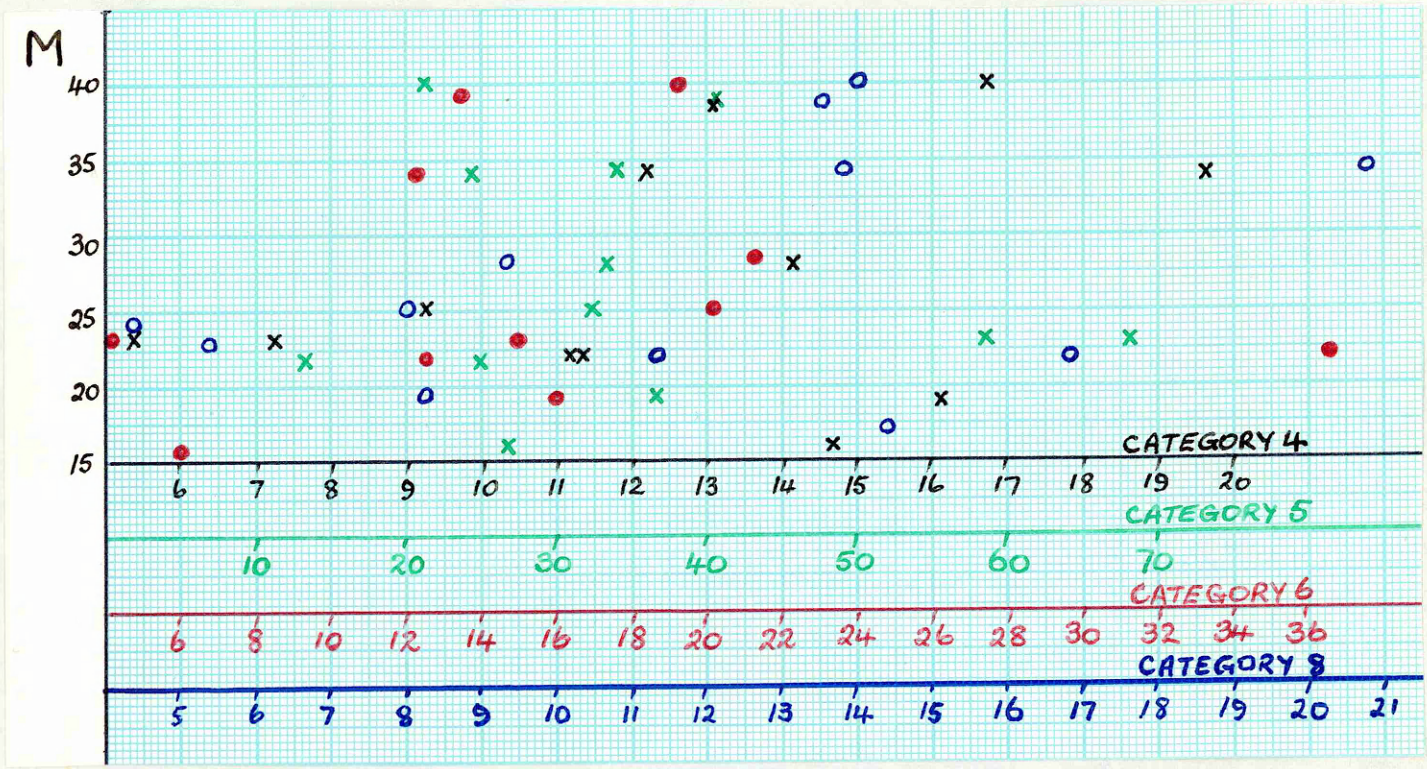


Table 13. product-moment correlations

F.I.A.C. %	4	5	6	8
Personality factor 'M'	.36 n.s.	-.13 n.s.	-.05 n.s.	.39 n.s.

Hypothesis 5 is rejected

5.7 Hypothesis 6

fig. 14. sub-sample of men

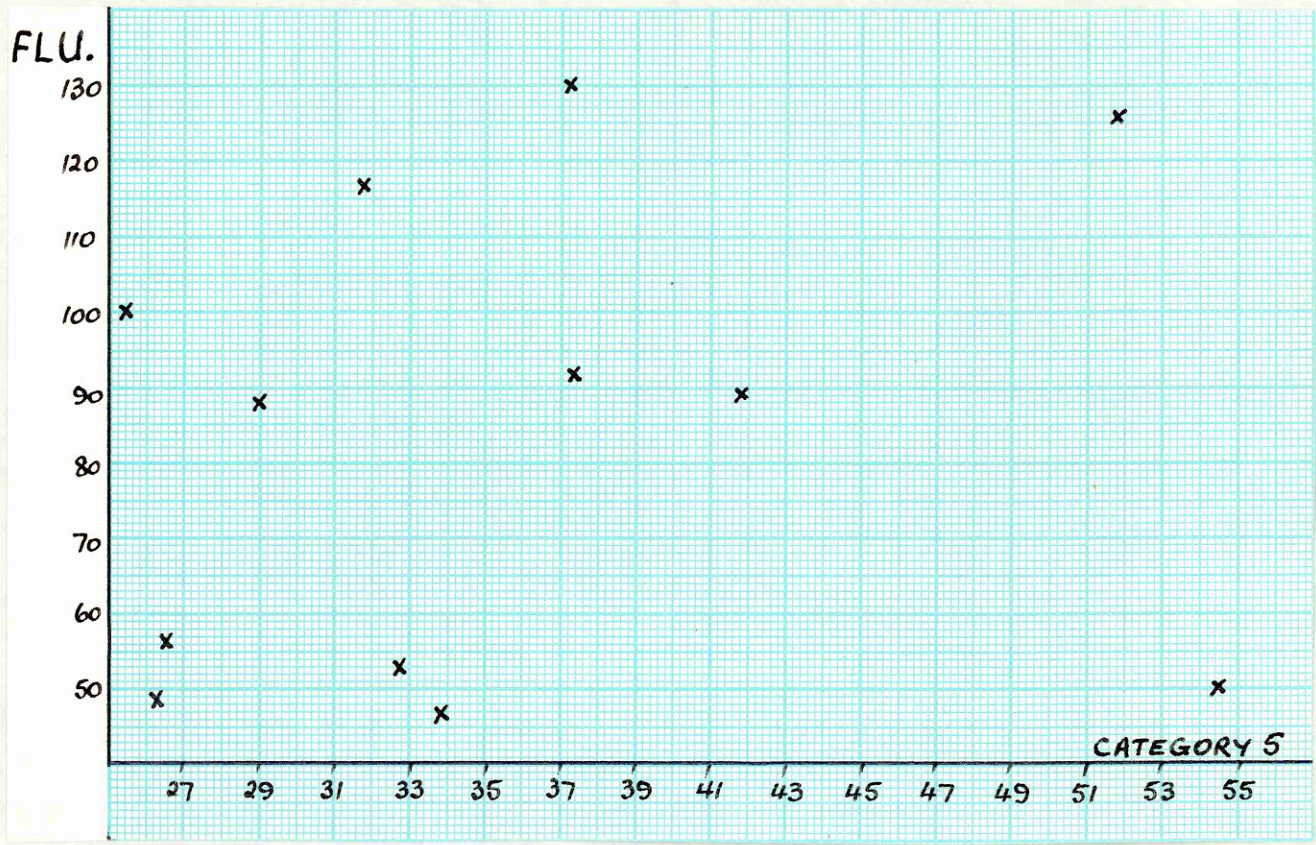


fig. 15. sub-sample of women

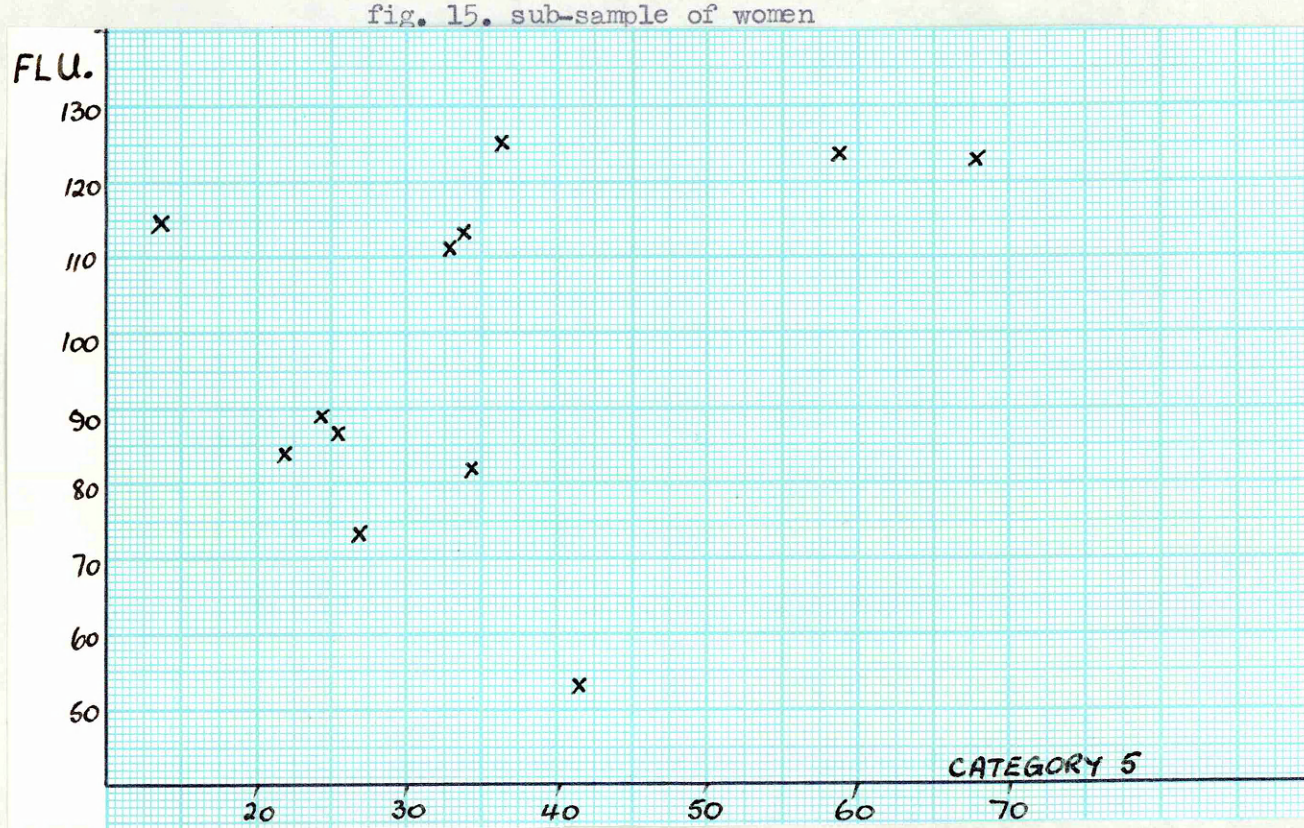


Table 14. product-moment correlations

F.I.A.C. %	5
Fluency	.18 n.s. (men)
	.35 n.s. (women)

Hypothesis 6 is rejected

5.8 Hypothesis 7

fig. 16. sub-sample of men

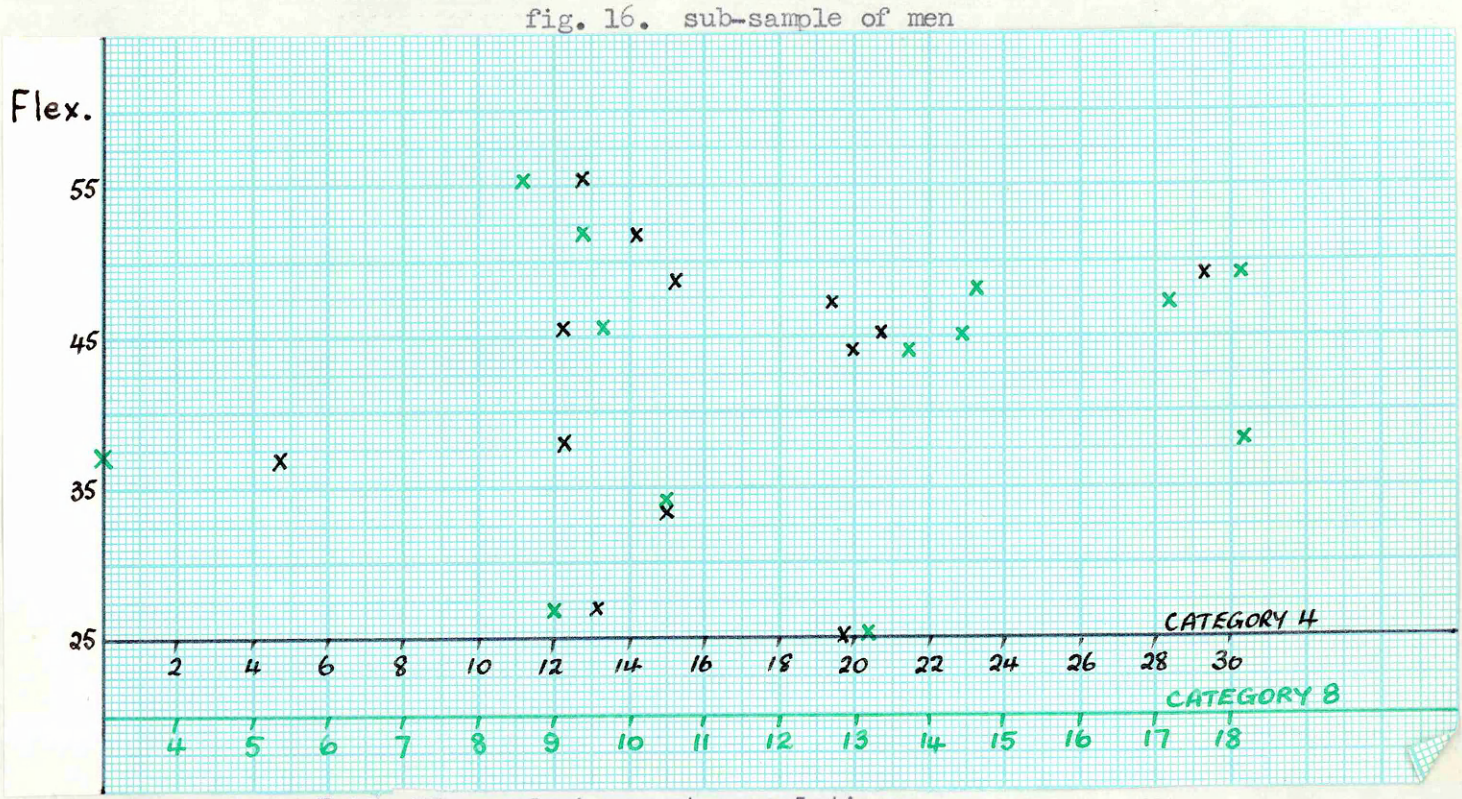


Table 15. product-moment correlations

F.I.A.C. %	4	8
Flexibility	.15 n.s.	.14 n.s.

fig. 17. sub-sample of women

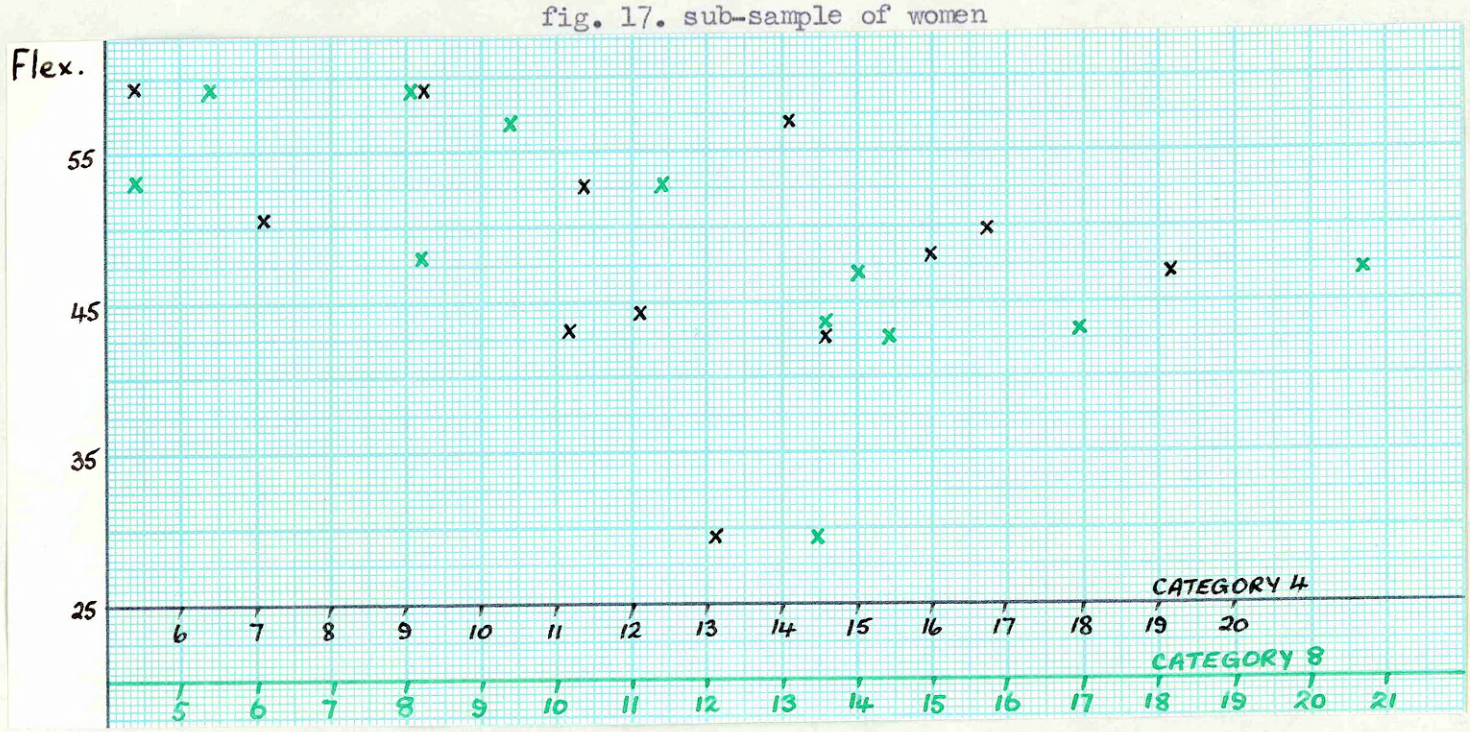


Table 16. product-moment correlations

F.I.A.C. %	4	8
Flexibility	-.43 n.s.	-.60 *

Hypothesis 7 is rejected

* significant at p.05 level

5.9 Several statistically significant coefficients were obtained for pairs of variables which had not been subjects of the research hypotheses. These are tabulated here for later discussion.

Table 17. sub-sample of men

Dependent var. F.I.A.C. %	1	3	5	8	9	10	TT	TQR	PIR
Indep.variable									
'C'	.70*								
'E'									.64
'L'			-.72*						
'M'					.57				
'N'	.58								
'O'						.55			
$'Q_1'$				-.74*			.64		
$'Q_2'$		-.58							
'Q4'	-.68*								
Anx.			-.62						
Flu.					-.56	-.69*			
Flex.						-.76*			
Orig.									-.57

*p.01; otherwise p.05

Table 18. sub-sample of women

Dependent var. F.I.A.C. %	2	4	5	6	8	9	10	TT	TRR	PIR
'A'				-.64			-.67			
'G'	.67				.61			-.62	.64	
'M'						-.58				-.83*
$'Q_3'$			-.60	-.70						
Flu.	-.61				-.72*				-.60	
Flex.					-.60					
Orig.	-.61	-.59			-.69			.61	-.66	

*p.01; otherwise p.05

CHAPTER 6

DISCUSSION

6.1 Origins of the hypotheses

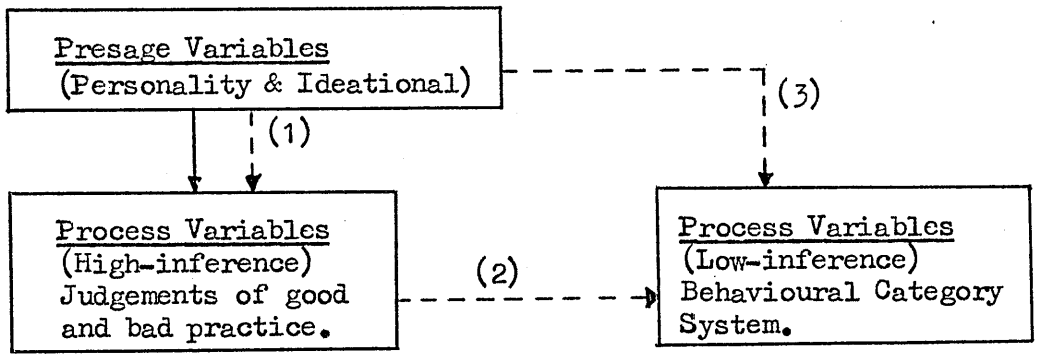
The hypotheses were formulated from a convergence of two approaches. Firstly, by a review of research linking personality and teaching success as variously defined, but especially in relation to classroom performance judged by experts. What information is available on the bases of such judgements indicates no broad consensus on the criteria by which competence is assessed. It was therefore necessary to examine the measured variables as potentially causal factors in the presage - process system and to speculate upon those behavioural acts (process) which would be likely to create a favourable impression in observers. That is, to reason why such results might have been obtained.

Secondly, in an attempt to instil a greater degree of objectivity into the observation of teaching, an examination of a low-inference coding system was made and relations between categorized acts and the presage variables were postulated on the basis of the internal logic of the personality, ideational and category systems reviewed. This latter stage was kept as free as possible from the evaluative perspective of "good" and "bad" practice which is the essence of the earlier step. The process is shown in fig. 18 in which the dotted lines indicate the directions of inquiry.

6.2 Hypothesis 1.

The independent variable, factor 'A' scores, relates to the Affectothymia - Sizothymia dimension whose positive pole

fig. 18



Paths of inquiry:

1. What were the approved/disapproved behaviours likely to be?
2. How would these relate to the category system?
3. What hypotheses may be formulated to link presage and process (low-inference) variables directly?

encompasses the attributes described by Cattell & Eber (1962), listed in section 2.62, and whose negative aspect stresses 'inflexibility, coldness, secretiveness, timidity, cynical hostility, and pessimism' (op.cit. and Cattell, 1946).

In relation to Category 1 statements, it was hypothesised that a positive correlation would be found between higher factor 'A' scores and the teachers' acceptance and clarification of pupils' attitudes and feelings. This is consistent with the 'sociability' of the affectothymic temperament which includes 'sympathetic conversational participation', that is, sociability is not to be understood here as merely an outgoing quality of the individual, but is also to be seen as an emotional receptivity; a sensitivity to the feelings of others.

The results indicate that the hypothesis is not tenable on the basis of the information gathered in this study. This is not held to be a disappointing outcome in view of the scoring constraints posed by this behavioural category and the difficulties experienced

by other researchers in interpreting statistical analyses of similar data. Wragg (1972) noted an extremely small proportion of tallies falling into category 1 and Morgan & Woerdehoff (1969) comment that their relative proportion to the total of tallies was so small that they "were lost" in multiple correlation procedures. The obtained percentages of category 1 tallies from the two sub-samples amounted to approximately .01, only 8 instances for men and 11 for women were recorded out of totals exceeding 8,000 for each sex. In addition, half the men and two-thirds of the women yielded no data in this category. This resulted in marked skewing of the distributions (see figs. 6 & 7) and illustrates the insecure foundations of the coefficients given in tables 7 & 8.

Flanders (1970) refers to category 1 behaviours as relatively rare and infrequent teacher statements, and he stresses the need to record such events whenever they occur in order to maximize information. Two pertinent questions at this point seem to be: why are such events rare, and what is their significance if they are so infrequent? The former is answerable in part by the protocol for scoring this category; the teacher must explicitly name or otherwise indicate the emotion or feeling before the observer can code "1". A partial explanation of both questions is that there is justification for a category of respondent teacher-talk distinct from 2 and 3 which is predominantly concerned with the affective. Flanders does not elaborate on the function of category 1 statements but they are clearly expressive in contrast to instrumental transactions, being concerned with the regulation of inter-personal relations and not the transmission

of lesson content. The infrequency of these behaviours can also be attributed to cultural effects via the roles of teacher and taught as traditionally perceived, where both parties tend to suppress positive and negative emotional reactions in the classroom.

In his analysis of teachers' roles, Hoyle (1969) comments that socialization is the most significant function of the teacher in the infants school and that the inculcation of 'acceptable' standards of social behaviour is of great importance there. Under these circumstances it might be expected that category 1 behaviours would assume a greater numerical and functional significance and, insofar as the Flanders System claims to have application across a wide age spectrum, the inclusion of this category is logical although not necessarily appropriate where older pupils are being taught.

The occurrence of statistically significant coefficients for factors 'C', 'N' and 'Q_4' merits some comment, qualified by the strong reservations expressed earlier about the basis of these correlations. Inspection of stepwise regression analysis summary (Appendix VI, table I (i)) shows that factor 'C', which accounts for a considerable portion of the criterion variance, has been selected at step 1 and factor 'Q_4' at step 9 in a sequence of 10 steps, while factor 'N' has not been selected at all. This indicates that variance represented by 'Q_4' and 'N' is substantially common either to 'C' or to the other selected variables, none of which separately approaches significance at the .05 level. If this correlation of factor 'C' were a credible effect it would suggest that the personal qualities of 'emotional

calmness' and 'realism about problems', among others, were determinants of category 1. This would not be inconsistent with the specified behaviours but there is every likelihood that the result is spurious.

6.3 "Nothing succeeds like success" is an aphorism which sums up in common sense terms the prevailing belief that reward is generally more effective, and perhaps more acceptable morally than is punishment, in the teacher's management of learning. Whether one takes a mechanistic or a cognitive perspective on behaviour, reinforcement or reward is likely to increase the incidence of approved acts through operant conditioning or improved self-esteem. Therefore the effectiveness of teaching judged by the extent to which lessons meet the teachers' objectives is related to the directing influence of the giving or witholding of praise.

Supervisors of practising student teachers are concerned with the quality of relationship established between teacher and taught as part of this general aura of effectiveness as is evidenced by the content of the types of teacher appraisal schedules reviewed by Norris (1975); and an index of this relationship is what Flanders calls the 'genuine praise' given by the teacher.

Genuine praise (Flanders, 1970, p.41) can be distinguished from the routine habitual usage of words like 'good' and 'right', which may have achieved redundancy as effective reinforcers, in that it "...often takes longer than three seconds to express." This assertion which tests genuineness by the exceeding of the observational quantum appears inconsistent with the description

of category 2 acts where nodding the head, saying "Um, hm" or "Go on" are included as legitimate exemplars. In tallying events in category 2 the present observer included all vocalized and otherwise signalled teacher approvals which occurred during a recording segment.

It was felt that there would be a close relationship between Affectothymia scores and rewarding behaviour which would be consonant with the factor description predisposing the individual to emotional receptivity and expressiveness. The results (tables 7 & 8) do shew positive correlations which however fall short of the .05 level of significance.

The summaries of the regression analyses for category 2 (Appendix VI, tables I(ii) and II(ii)) show different patterns for men and women. For the male sub-sample factor 'A' is first selected, makes by far the largest single contribution to criterion variance, and the proportion of the variance accounted for by it in successive steps increases steadily to indicate the relatively greater importance of Affectothymia when factors are considered together. The result for women gives a different ordering of traits. Although factor 'G' has the highest zero order r of 0.67 and is the first selected variable, its relative importance diminishes as factors 'N' and 'F' enter the running. These three make individual contributions to multiple R of about the same order and taken together account for a very large part of the criterion variance. The absolute size of each factor's contribution to multiple R is not of interest in this or in the other multiple regression analyses and the data in their present form are not intended to shew this. R is known to capitalize on

correlated errors and the smaller the sample in relation to the number of variables involved, the greater is the exaggeration of the size of R. In the present work the samples are small and it is the relative amounts of variance attributable to each factor that are of interest.

When the trends in the regression analyses are interpretated in terms of the factor descriptions there is a noticeable tendency for Affectothymia in men to contribute toward praising behaviour. The same category in women is associated with being easy to please, artless as opposed to shrewd ('N' -); being attentive to people, conscientious, responsible, emotionally stable ('G'); and being subdued, introspective, with a tendency toward the obsessional ('F' -). The responsibility represented in factor 'G' is not just a rational politeness but a 'categorical imperative' (Cattell, 1965). This goes along with a sense of duty or obligation; one might say vocation, and might be connected with a particular perception of what 'good' teachers do, viz. give praise. This is linked with the sober qualities of quiet persistence and dedication of 'F' - and the directly relevant quality of being easy to please ('N'-).

Some justification for the trend linking Affectothymia in the men with category 2 can be made which parallels the findings of several researchers, e.g. Montross (1954); Henjum (1969), that there is evidence that higher scorers on factor 'A' are rated as more competent teachers. The trend for factor 'G' scores in women to contribute to the same kind of behaviours also finds ground in similar reports by, for example, Warburton et al (1963); Tarpey (1965); Henjum (1969); and Davis & Satterly (1969).

The occurrence of dissimilar, although not incompatible, findings for men and women in the present study bears closer examination. Some previous reports have suggested that men and women teachers achieve success through different personality traits appropriate to their sex, e.g. Chabassol (1968); McClain (1968); Gough et al (1968); but there may well be other less fundamental causes for the present observed differences in the experimental data.

The low correlation of factor 'A' with category 2 for women may be due to the restriction of range in the predictor measure. The array in fig. 8 reveals over 80% of the points plotted lying between scores of 18 and 20 on factor 'A', with several tied scores, while the comparable data for men cover twelve points on the same scale. It is therefore probable that the result for women is a product of this sampling accident with respect to factor 'A'.

Scores on Fluency and Originality also produced significant and almost identical coefficients for category 2 among the women, although the regression analysis selected only Originality (at step 3), suggesting that substantial communality exists between these two factors. The proportionate contribution of Originality to criterion variance is quite large, amounting to over half that made by each of the principal correlates 'N'-, 'G' and 'F-', and indicates the tendency for the capacity for original thinking in the teacher to be negatively related to the awarding of praise. This is difficult to explain in terms of this criterion considered alone but perusal of table 18 shows that nearly identical coefficients in the same direction were found for Fluency and

Originality in relation to the Teacher Response Ratio (TRR) and that Originality is positively and significantly related to the total Teacher Talk (TT); Fluency at r=0.57 just falling short of the significance threshold of 0.58.

When considered in this wider context, the discernible tendency is for divergent thinking capacities to relate to greater teacher participation in the verbal interchanges but for less of the teacher talk to be in response to pupils' contributions. This is supported by the data in Appendix VI, tables 1(xi), 1(xiv), 2(xi) and 2(xii), where ideational factors make the largest single contribution in each case to Teacher Talk (both sexes), Pupil Initiation Ratio (PIR) (men), and Teacher Response Ration (women).

Although not entirely consistent, these outcomes taken together suggest that the capacity to generate a flow and a variety of ideas predisposes the teacher to dominate the verbal interaction at the cost of pupil initiatives and the teacher's own respondent discourse.

6.4 The substantial positive although non-significant coefficients obtained for Affectothymia and category 3 reflects the decreasing affective connotations of the respondent behaviour represented in categories 1 to 3. Flanders (1970) claims that category 3 is associated with higher classroom measures of both content achievement and positive pupil attitudes toward schoolwork and the teacher, and represents the teacher's cognitive orientation in accepting and incorporating pupils expressed ideas into the lesson. Insofar as the sensitivity and warmth toward others may play a part in such acceptance, factor 'A' makes only a minor contribution as is shown in the regression analyses (Appendix VI,

tables 1(iii) and 2(iii)) where 'A' is selected very late for inclusion in the computer runs for both sub-samples. These findings run counter to those by Wragg (1972) who found factor 'A' to correlate 0.31 with category 3 on a much larger sample.

The negative coefficient for factor 'Q_2' just reaches significance for men and the coefficient for women is small but in the same direction. The regression analyses reveal a very small involvement of this factor in category 3 behaviour, 'Q_2' being selected at step 7 for women and not at all for men. This indicates that the factor scores contain negligible predictive information that is not available via the other selected variables. What minute influence this factor may have does appear to be in the right direction, higher scores 'prefer their own decisions and discount the opinions of others', a set of qualities hardly likely to predispose a teacher to value his pupils' contributions.

6.5 Category 9 (pupil initiation), may be seen in part as a logical complement of category 3 (teachers' accepting behaviour), as was reasoned earlier (at 3.7) that there would be a causal link between the incidence of accepting behaviours as reinforcers and pupil initiatives as operants, and that both categories would be correlates of Affectothymia scores. The evidence provided by the coefficients and the multiple regressions lends no support for this hypothesised connexion between factor 'A' and category 9. For both sub-samples r is very small and factor 'A' was not selected at any step in the analyses.

These outcomes are understandable in the same terms outlined in the discussion of category 3 and factor 'A', but there is a particular problem relating to the scoring of category 9 which must have contributed to the result obtained. Pupil initiatives can take many forms which satisfy the protocols 'expressing own ideas' and 'going beyond the existing structure', and there is an obvious dichotomy in acts of initiation which are broadly in line with the proper business of the lesson and those which are, for example, red herrings, facetious comments, incipient rebellion, etc. The acceptance of all pupil initiatives within this category must confound behaviours which are not even close pedagogical equivalents and make valid interpretations of the data difficult to achieve. In view of this the discussion of other correlates of category 9 which follows needs to be weighed against the crudeness of the data which the tallies represent.

Perhaps symptomatic of this problem is the finding that factor 'M' makes the largest contribution to the variance in both sub-samples yet appears to do so in opposite directions: higher scoring men and lower scoring women registering more tallies. This is apparent in the regression analyses for category 9 and the trend is also marked in the derived measures (PIRs).

The findings of teacher competence associated with factor 'M' reviewed earlier shew an ambivalence which may represent basic sex differences as suggested by Chabassol (1968) since the majority of investigators failed to analyse their data for each sex separately. It is also unclear in what manner pupil initiatives would be judged to indicate teacher competence; it might be thought more relevant

to observe how the teacher deals with initiatives whether these are broadly lesson-oriented or off-target, but there is no evidence among the teachers' respondent behaviour data for a substantial role for factor 'M'. This suggests that many of the initiatives were not of the lesson-oriented type and some indirect confirmation of this may be had from the comparatively greater involvement of factor 'M', in the predicted directions, with category 7 (teacher criticizes; justifies authority, etc.). Additionally, higher scores on factor 'E' for men, indicating austerity, extrapunitiveness and authoritarianism, correlate with pupil initiatives and account for about a fifth of the observed variance in both sub-samples. However, these interpretations remain highly speculative and are, at best, held only very tentatively.

The coefficient for Fluency with category 9 reaches the .05 level of significance for men and, whilst the coefficient for women is very small, ideational factors generally account for between 20 and 30% of the observed variance in both sub-samples and are another index of the way in which divergent thinking capacity in the teacher seems to go with a reduction in time available for registrable pupil initiatives.

6.6 Categories 1 - 3 and 9 form a logical cluster of pupil initiative and teacher respondent behaviours, and before going on to consider results linked with other categories, some summary of the trends discussed is appropriate here.

The hypothesis proposed positive links between Affectothymic temperament in the teacher, teacher respondent behaviours, and pupil initiatives. No statistically significant correlations

uphold the hypothesis although there are trends in the regression analyses which indicate that the more Affectothymic men gave more praise. A similar finding for women did not occur, possibly as a result of a statistical artifact. Difficulty was experienced in drawing any valid conclusions from category 1 data owing to the extreme rarity of these behaviours. Doubts were raised about the suitability of this and other categories for several reasons. Statistically significant findings for other personality factors in relation to these categories were discussed in the context of previously reported claims and the probable ways in which teacher behaviours would be judged. Ideational factors showed marked similarities in their relationships with the Flanders Categories and there was a noticeable tendency for the teachers' divergent thinking capacities to correlate positively with all forms of teacher talk and therefore negatively with pupil initiatives.

6.7 Hypothesis 2

This is accepted without much reservation. Many studies reviewed have produced conflicting or nonsignificant findings for the relationship between intelligence and teaching success. The results in Appendix V show generally small coefficients for factor 'B' with most categories. The regression analyses selected this factor in only half of the twenty-eight computer runs and then at positions ranging from third to ninth with a median step position of seventh. The finding is consistent with other work that has used full scale intelligence tests and it supports the idea that once selection for teacher education has been made, students are relatively homogeneous on this trait when compared with the general population. This is confirmed by the statistics

for the distribution of factor 'B' scores in the sameple and in the general population.

	Mean	S.D.
General U.K. Population (1970)* forms A + B	14.41	3.47
Sample " "	17.40	2.99

Some reference needs to be made to the rather restricted notion of intelligence exemplified in the questionnaire items and the small number of questions which load on factor 'B' when comparisons are made with a full scale test. Out of the thirteen items measuring 'B' in each form of the 16 P.F.Q., eleven are verbal; classifications and analogies; and two are numerical/spatial, yet require verbal reasoning. This underlines Vernon's (1963) concern that an instrument which claims to yield measures on so many different dimensions can do so only at the expense of reliability since the number of items which refer to each factor must be severely restricted if the length of the questionnaire is to be held within manageable limits. It is felt that had intelligence been a major concern in the present work the use of a full scale test, of which there is a wide variety, would have been imperative.

6.8 Hypothesis 3

The proposed link between factor 'G' scores and the Flanders Categories was made with less confidence than was possible in the case of factor 'A', although the evidence concerning 'conscientiousness' or 'superego strength' and teacher competence is less equivocal,

---

*from Saville P. (1972). The British Standardisation of the 16 P.F. - Supplement of Norms, Slough, N.F.E.R. Publishing Co.

i.e. none of the research reports reviewed indicated a negative relationship. At 3.7 it was hypothesised that the dominance by a sense of duty (elaborated by Cattell 1946 and 1965) as persevering, determined, insistently ordered, exacting planful, and being driven by the "inner categorical imperative of their essential superego") would predispose higher 'G' scoring students to careful lesson planning and meticulous attention to the host of other teaching-related activities which take place outside lessons. One could imagine, and perhaps name from personal experience, those students whose behaviour follows the pattern of careful lesson planning in terms of proposed techniques and ordering of resources, coherently reported in a neat file; who are eager to seek and accept the supervisor's advice, who 'fill the unforgiving minute' and make suitably modest evaluations of their own efforts.

It is easy to understand how these behaviours could not fail to impress a supervisor as to the students' 'attitude' and 'commitment' to the global aspects of teaching. That these acts are perceived as impressive, at least by students themselves, has been indicated by Sorenson (1967), who asked a large sample what they would tell their best friends to do in order to achieve a top grade from their supervisors. From his results, 50% said "be well organised", 40% said "follow advice without question", and 19% said "cultivate him"(ask advice when you don't need it*).

Cynicism apart, this catalogue of advice admirably shows that the present assessment of the susceptibilities of supervisors may have some foundation and indicates the importance of non-classroom events in the judgements of student teachers.

The hypothesis relating to F.I.A.C. rests on less secure bases and it was reasoned that those same qualities of planfulness, etc., characteristic of higher 'G' scores, would make for inflexibility, a lack of spontaneity, and hence more teacher directed interaction. This general hypothesis is unacceptable although some details merit comment.

Categories 4 and 8 comprise all the teacher questions - pupils answer sequences of behaviours and the obtained coefficients for both these with factor 'G' show a close approximation in size which underlines the cause and effect nature of their relationship (see tables 9 and 10), although only r for category 8 (women) achieves statistical significance while r for category 4 for the same sub-sample approaches it at 0.51. Despite the magnitude of these coefficients the regression analyses indicate a negligible involvement of factor 'G' among the other predictors of both categories. Only in one case - category 4 for women - was 'G' selected, at step 4, and its contribution to net variance reached about 7% when R was maximum. Similarly in the case of the derived TQR, 'G' was not selected for men and appeared only at step 9 for the women's group and made only a minute contribution there.

There are problems with category 4, whose nature have been touched upon in the review of literature and in connexion with other categories in this discussion, namely the grossness of the data which here results from lumping together all teachers' questions. Flanders (1970) is aware of this shortcoming and the difficulties it creates in decoding and interpreting interaction records. He distinguishes several different functions of questioning and elaborates on various types, viz: broad vs. narrow, and open vs. closed, but these are also open to misinterpretation as MacLeod (1975)

has indicated.

It is felt here that a category as crude as 'teacher asks questions', even when questions intended to be rhetorical or critical are excluded, is too broad to discriminate between what are seen as pivotal points in a lesson. The mere quantification of questioning behaviour is bound to overlook not only subtle shifts of emphasis which bear upon the direction of the lesson, but also the differences between, for example, a spate of superficial queries which harp upon the same theme, and a few thoughtful probings which lift the level of inquiry to an altogether different plane of activity.

There is no evidence to support the hypothesis that the qualities associated with factor 'G' contribute anything significant from the pool of predictors towards the incidence of teacher questioning/pupil respondent behaviours as recorded in these categories.

Similarly, the idea that the qualities of conscientiousness, planfulness, etc. in higher 'G' scores would predispose students in their keenness to lecture, to 'tell all', to instruct, is not supportable. Factor 'G' was unselected in the regression analyses for categories 5 and 6 for men and, although included in both for women, declined in relative importance during the computations and ended up making contributions of less than 1% and less than 10% respectively to net variance in category 5 and 6 behaviours.

Additionally there are problems attaching to the measurement of factor 'G'. The questionnaire data on this factor with its connotations of conscientiousness etc., is likely to suffer considerably from motivational distortion along the lines described by Thorndike (1971) discussed at 2.50. An examination of the items loading on 'G' in both

forms of the 16 P.F.Q. shows several areas thought to be particularly sensitive to this effect when the nature of the sample is taken into consideration. A few examples will suffice:

Item content	occurrence
(a) care for other individuals and their property	4
(b) dislike of disorder/untidiness/ carelessness	4
(c) value of following rules/behaving morally	5
(d) service to the community	2
(e) value of planning/avoiding waste of time	1
(f) liking work needing conscientiousness and precision	1

This leads one to the conclusion that the factor 'G' data and the correlations derived from them may be less valid than those obtained in other areas of the investigation.

6.9 Hypothesis 4

Lamke (1951), Tarpey (1965), and Henjum (1969) report positive associations between higher scores on factor 'H' and teaching marks. A brief consideration of the attributes of low scores provides grounds for reasoning why this might be so. The factor description provided by Cattell & Eber (1962) characterises the Threctic (H-) individual as restrained, cautious, tending to be slow and impeded in speech and expression, usually having feelings of inferiority, and not being given to keeping in touch with all that is going on around him. These would seem to provide a recipe for failure for any teacher in the classroom. Conversely the

imaginativeness, spontaneity and autonomic toughness (stamina) of Parmic ('H'+) subjects would be seen as desirable qualities, and a corollary of Parmia is a tendency towards oral fluency which may also be an impressive feature of a teacher's behaviour. Cattell (1965) reports that Parmic subjects often display this oral facility to excess, and he cites the outcomes of researches into behaviour in small groups where high 'H' scoring individuals tended to ramble on and to provoke criticism from their more laconic colleagues.

It was hypothesised here that adventurous boldness, insusceptibility to inhibition, and oral fluency of higher 'H' scorers would have correlates in Flanders Category 5 (lecturing) behaviour. The results give little support to this proposition. Factor 'H' did not feature in the regression analysis for women and, although selected in second place for men, it played a reducing role with successive steps and finally contributed about 10% to net criterion variance. If we consider Teacher Talk as a whole there is no sure indication that higher 'H' scores have gone with oral fluency and a survey of the separate categories 1 to 7 reveals no firm evidence for a link in the direction proposed. As the Flanders Category System is unquestionably sensitive to quantities of talk, and would have registered any tendency toward garrulity, it must be assumed either that those other qualities associated with higher 'H' scores have impressed observers of student teachers who judgements form the bases of the researches reviewed, or that the propensity forming the basis of the present hypothesis has not been realised in this sample.

6.10 Hypothesis 5

Categories 4, 5 and 6, and the pupil answer category 8 do not

relate to factor 'M' scores in any consistent way. The ambivalence discussed at 6.5 in respect of the relationship between this factor and category 9 is again represented in the contrary polarities of the coefficients shown in tables 12 and 13. There is also a probability that the qualities described by factor 'M' do not relate in any conceptually linear way with what might be judged to be good teaching. Low scores on 'M' (Praxernia), denote an attention to practical matters, concern over detail, conventionality, anxiety to 'do the right thing', and unimaginativeness. These properties, depending upon their intensity, may be perceived as either positive or negative indicators of competence, as might their opposites associated with high 'M' scores: imaginatively-creative, self-motivated, unconventional Bohemian. Higher scores on 'M' also predispose individuals to adopt unrealistic positions on emotional issues, to perceive reality falsely in accordance with their own wishes, and to be disinclined to alter their ideas in response to brute facts.

The regression analyses selected factor 'M' in only two instances: categories 5 & 6 for men. In each case the stepwise procedure accorded less importance to 'M' as the analysis continued and only relatively small amounts of variance were 'explained' by this factor.

6.11 Hypothesis 6

The types of behaviours accommodated in category 5 include lecturing, expressing opinions, interjecting thoughts, and offhand comments which, totalled up, account for a very large proportion of the verbal interaction between teacher and taught.

This telling activity usually produces the highest frequency of tallies (ranging up to 67.9% in this study) and hence a greater stability in the codings in comparison with other categories.

It was reasoned that a greater capacity for ideational fluency would result in more teacher domination of the interaction via the kind of behaviours just outlined; the effect being a more expository style of teaching revealed by an increased proportion of tallies in category 5.

Neither the zero order correlations for Fluency nor the regression analyses confirm this proposal. Table 14 shows Fluency to correlate positively with lecturing for both subsamples although the coefficients are small, and the summaries in Appendix VI indicate that only for men is this factor selected in the analysis of prediction of category 5, and then it makes a mere 10% contribution to net variance. On the other hand, the total Teacher Talk for men is substantially dependent upon Fluency, which together with Originality, accounts for over half the variance in this criterion. In the women's case there is little to parallel these trends in the same categories, but Fluency does make a major contribution to the '4-8' pattern of interaction.

Herein lies the limitation of treating the separate categories as independent behaviours, since although they are themselves static, they represent a dynamic activity whose events have strong causal links. The capacity to generate a greater flow of ideas need not necessarily be realised and, if it does find expression, it may be through modes of verbalisation appropriate to the moment and which transcend individual categories.

Their effects would consequently be latent but irretrievable in the data expressed as proportions of time spent on individual classes of activities.

The findings by Morgan & Woerdehoff (1969) that lecturing correlated positively with ideational fluency was not replicable in this study although there are indications of a tendency for Fluency scores to appear prominently in multiple regressions for categories describing teacher initiatives.

Knoell (1953), and Montross (1954) both found relationships between rated teacher competence and fluency scores, and Crocker (1968) comments upon the likelihood that supervisors are impressed by articulate students. The present results cannot support the notion that such a quality is reflected solely in telling behaviours.

6.12 Hypothesis 7

Crocker (1968) found Flexibility to be the most consistent predictor of rated teacher success and claims for a concept of flexibility having similarly positive links with competence have been made by Joyce & Hodges (1966), and Sprinthall (1966). Reservations about this apparent consensus were made at 2.85 when a contrast was pointed between Flexibility in the cognitive capacity sense and flexibility within the meaning of adaptability to circumstances, which would find ground in personality variables. This latter manifestation could be seen as a positive attribute in students who have to cope with the wide range of adjustments needed to reconcile demands made upon them and to 'fit in' with teaching practice schools as Ryan (1966) suggests. The ideational capacity, on the other hand, was considered as a predictor of the number of different approaches a teacher might make to a subject

under discussion, and it was hypothesised that greater Flexibility would correlate with a higher incidence of questioning resulting from this diversity of approach. Such a finding was made by Morgan & Woerdehoff (1969). The serious misgivings about category 4, data expressed in 6.8 again apply but with added emphasis because the terms of the hypothesis make the assumption that more questioning necessarily indicates that different types of questions are being asked and this is not logically tenable. In the event no significant coefficients were obtained which related Flexibility to questioning or to pupil respondent behaviour. Contributions to criteria variance in the regression analyses were either very small or negative.

6.13 The discussion of the hypothesised correlates of categories 4, 5, 6 and 8 has proceeded so far without detailed reference being made to those statistically significant outcomes which were not expected. Tables 17 and 18 show these findings which will now be considered to see to what extent they can add psychological meaning to the foregoing treatment, and what other factors appear to contribute substantially to the variance in categories 4 to 8.

Originality has the highest correlation with category 4 for women (sigt. at 0.5 level) and makes by far the largest contribution to variance there in the regression analysis. One interpretation is that low Originality goes with a higher incidence of questioning. This cannot be accepted as theoretically consistent with the concept of Originality defined by the Torrance Tests unless the type of questioning was routine, superficial and limited in scope, and this is a qualitative distinction that cannot be made.

All coefficients for category 4 in men fall short of the 5% confidence threshold but the second-order factor 'Anxiety' and factor 'Q_1' together account for over half the variance. This is not inconsistent with Birkin's (1971) finding that the more anxious teachers tended to talk more, although in this present instance the effect seems to be restricted to questioning; the combination of higher 'Anxiety' and lower scores on 'Q_1' indicating that individuals who are ill-at-ease, who respect established ideas and are confident in what they are taught to believe, tend to adopt a questioning approach in their teaching.

Conversely, there is a suggestion that 'Anxiety' is negatively related to lecturing behaviour. This is shown by the obtained coefficients for factors 'L' and 'Anxiety' in table 17 and the increasing prominence of 'Q_4' in successive regression analyses; both 'L' and 'Q_4' making substantial contributions towards the second-order factor. A similar finding for women does not appear in identical terms but there is complementary evidence provided by the high coefficient for category 5 and factor 'Q_3' scores and, although the influence of this trait declined in the stepwise direction, 'Q_3' is also a component of factor 'Anxiety'.

Primary factor contributors towards 'Anxiety' appear to act differentially in respect of these two classifications of teacher talk, and behaviours coded in category 6 also find a significant correlate in factor 'Q_3' among the women students. This outcome has no corresponding result among the men but it may add to the weight of evidence against the notion that higher anxiety is a promoter of teacher talkativeness as such.

6.14 This discussion has been punctuated by criticism of the measurement techniques used in the research, some foreseen and referred to in general terms during earlier stages of the report, while others have become apparent in the course of the experimental work and in reflecting upon the results. These results themselves do not go very far toward clarifying the issues embodied in the original hypotheses and, apart from the criticisms already made, there are others which concern the research at both tactical and strategic levels which need to be voiced.

An observational approach using Flanders Interaction Analysis was chosen to improve objectivity, that is, the reliability of recording classroom events. This can be seen to have had undesirable effects in that, while permitting the minutiae of verbal interchange to be counted, it cannot differentiate qualitative differences among events which may, by their significance to the participants, override and invalidate data which ~~is~~ are merely quantitative. In the same way the elaborations of the Flanders technique in which sequential coding of events is designed to impart some directionality to the record cannot cope with this basic objection. The system assumes a priori a universal significance of classroom events and a functional equivalence of the behaviours coded within each category.

These difficulties pose the greatest problems in research like the present study which seeks correlates of data assumed to be relatively 'pure' but which are likely to be seriously confounded. All similar studies e.g. Morgan & Woerdehoff (1969), Birkin (1971), are most certainly affected and it is against this uncertainty that no great value is placed upon the finding in this work that there was no statistically significant disparity among the proportions of time allocated to the Flanders categories when the students were observed on several occasions. This is an outcome which replicates

Morgan & Woerdehoff's result but it may mask the occurrence of pedagogically significant differences to which the instrument is insensitive. This is not an argument against the application of Flanders' Analysis and other similar instruments in what may be described as legitimate contexts as, for instance, within training schemes where other variables are carefully controlled and specific behavioural changes in quantitative terms are sought, as has been described by Brown (1975) and Kleinberg (1975); but their use in less well defined situations can be seen to have serious weaknesses.

The computational technique and the number of variables used in this work have probably thrown up several spurious correlates which are indistinguishable from real effects. MacLeod (1975) comments upon studies employing large numbers of independent and dependent variables, where the statistical approach used is one relying upon probabilistic statements of significance, to emphasise the possible appearance of significant relationships explicable solely by the 'laws' of probability. Here the variables have yielded 210 zero order coefficients for each sub-sample, and there is thus a likelihood that several coefficients have reached the .05 level of statistical significance on these probabilistic grounds alone.

Different findings for the men's and women's sub-samples may owe as much to this effect as to the fundamental causes suggested by Chabassol (1968); in fact the provision of two separate groupings creates the opportunity for false as well as meaningful differences to come to light.

This study has examined an area of concern already well researched in terms of the amount of literature published on teacher

behaviour but its results provide no clear indication that the investigation of presage variables is as fruitful an avenue to explore as Birkin (1971) claims. Such a claim would need the supporting evidence of very large scale research with specified criteria for teacher behaviours more rigorously linked to the attainment of clearly articulated educational objectives. The apparent objectivity achieved by the application of numerical procedures in the absence of such process-product linkages is an illusion.

CHAPTER 7

SUMMARY OF CONCLUSIONS

1. Results for the group of men students showed Affectothymia to be a correlate of category 2 (teacher gives praise); a related finding for women linked this behaviour with the qualities of being easy to please, attentive to people, and conscientious ('N' and 'G').
2. Intelligence as measured by factor 'B' scores did not relate significantly to any of the behavioural categories.
3. There was a tendency for Ideational Factors to relate to higher incidences of teacher-initiated interaction.
4. There were indications that Fluency, Flexibility and Originality have a substantial communality of variance shown by the similarity of their correlation coefficients with all the behavioural categories.
5. Anxiety, either as a second-order factor or as represented by constituent source traits, tended to go with a greater amount of time spent asking questions.
6. No significant differences were found among the proportions of time allocated to the various Flanders Categories across several occasions of observation. This does not mean that changes of a qualitative nature did not occur.
7. The author has serious misgivings about the validity of some of the data collected by Flanders Interaction Analysis when used in research designs similar to the one employed here.

Bibliography

ADELMAN, C. & WALKER, R. (1975) Developing pictures for other frames: action research and case study, in Chanan & Delamont (1975).

ALLPORT, G.W. (1937) Personality: a psychological interpretation, New York: Holt.

ALLPORT, G.W. (1961) Pattern and growth in personality, New York: Holt, Rinehart & Winston.

ALLPORT, G.W. (1966) Traits revisited, American Psychologist, 21, 1-10.

ALLPORT, G.W. & ODBERT, H.S. (1936 Trait names: a psycho-lexical study, Psychol. Monogr. 47, (no.211), 1-171.

AMIDON, E.J. & HOUGH, J.B. (eds) (1967) Interaction Analysis, Reading, Massachusetts: Addison-Wesley.

ANDERSON, H.H. (Ed.) (1959) Creativity and its cultivation, New York: Harper.

ANNETT, J. (1969) Feedback and Human Behaviour, Harmondsworth: Penguin Books.

ARNOLD, J.E. (1962) Education for innovation, in Parnes, S.J. and Harding, H.F. (Eds.) A source book for creative thinking. New York: Scribner.

AUSTAD, C.A. (1972) Personality correlates of teacher performance in a microteaching laboratory, J. Exp. Educ. 40, (3), 1-5.

BANDURA, A. & WALTERS, R.H. (1963) Social learning and personality development, New York: Holt, Rinehart & Winston.

BARR, A.S. (1948) The measurement and prediction of teaching efficiency - a summary of investigations, J. Exp. Educ. 16, 203-283.

BARR, A.S. (1958) Problems associated with the measurement and prediction of teacher success, J. Ed. Res. 51, 695-699.

BECK, W.R. (1967) Pupils' perceptions of teacher-merit - factor analysis of five potential dimensions, J. Ed. Res. 61, (3), p.72.

BECKER, W.C. (1960) The matching of behaviour rating and questionnaire personality factors, Psychol. Bull. 57, 201-212.

BELLACK, A.A. et al (1966) The Language of the Classroom, New York: Teachers' College Press, Columbia Univ.

BIDDLE, B.J. & ELLENA, W.J. (1964) Contemporary research on teacher effectiveness, New York: Holt, Rinehart & Winston.

BIRKIN, T.A. (1971) Toward a model of instructional processes, - In Westbury, I. & Bellack, A.A. (Eds.) Research into classroom processes New York: Teachers College Press.

BRODY, N. (1972) Personality research and theory, London: Academic Press.

BROWN, G.A. (1975) 'Microteaching: research and developments', in Chanan & Delamont (1975).

BURT, C. (1949) The structure of the mind, B.J.Educ.Psychol., 19.

BURT, C. (1962) Critical notice: the psychology of creative ability, B.J.Educ.Psychol., 32, 292-8.

CARLILE, A.B. (1954) Predicting performance in the teaching profession, J. Educ. Res. 47, 641-668

CARR, H.A. & KINGSBURY, F.A. (1938) The concept of traits, Psychol. Rev. 45, 497-524.

CASHDAN, A. and
WHITEHEAD, J. (1971) Personality Growth and Learning: a source book, Open University Press.

CATTELL, R.B., (1946) Description and measurement of personality, New York: World Book Co.

CATTELL, R.B. (1950) Personality: a systematic, theoretical and factual study, New York: McGraw-Hill.

CATTELL, R.B. (1956) Personality and motivation theory based on structural measurement - In McCary, J.L. (Ed.) Psychology of personality, New York: Grove Press.

CATTELL, R.B. (1957) Personality and motivation structure and measurement, New York: World Book Co.

CATTELL, R.B. (1961) Theory of situational, instrumental, second-order, and refraction factors in personality structure research. Psychological Bulletin, 58, (2), pp.160-174.

CATTELL, R. B. (1965) The scientific analysis of personality, Harmondsworth: Penguin Books.

CATTELL, R.B. (Ed.) (1966) Handbook of multivariate experimental psychology, Chicago: Rand McNally.

CATTELL, R.B. & BUTCHER, H.J. (1968) The prediction of achievement and creativity, Indianapolis: Bobbs-Merrill.

CATTELL, R.B. & EBER, H.W. (1962) Manual for the 16 P.F. questionnaire, Illinois, I.P.A.T.

CATTELL, R.B. & EBER, H.W. (1967) Manual for forms A and B, 16 Personality Factor Questionnaire, Illinois, I.P.A.T.

CATTELL, R.B. & SAUNDERS, D.R. (1950) Interrelation and matching of personality factors from behavior rating, questionnaire and objective test data - Reported in Cattell & Butcher (1968).

CHABASSOL, D.J. (1968) The possession of certain attitudes as predictors of success in practical teaching, J. Educ. Res. 61, (7), 304-306.

CHANAN, G. & DELAMONT, S. (1975) Frontiers of classroom research, Slough: N.F.E.R.

COLLINS, M. (1958) Some correlates of inadequate teaching ability, Ph.D. thesis, Univ. of London.

CORTIS, G.A. (1968) Predicting student performance in colleges of education, B.J. Educ. Psychol. 38, (2), 115-122.

CROCKER, A.C. (1968) Predicting teaching success, Educ. for Teaching, 76, 49-52.

CROCKER, A.C. (1973) Student teaching success, Ph.D. thesis, Univ. of Durham.

CRONBACH, L.J. (1960) Essentials of psychological testing, New York: Harper (2nd Edition).

DAVIS, T. & SATTERLY, D.J. (1969) Personality profiles of student teachers, B.J. Educ. Psychol. 39, (2), 183-187.

DELAMONT, S. (1975) Introduction to Chanan, G. & Delamont, S. (eds) (1975) Frontiers of Classroom Research, Slough: N.F.E.R.

DOMAS, S.J. & TIEDEMAN, D.V. (1950) Teacher competence: an annotated bibliography, J. Exp. Ed. XIX, (2) 101-218.

DREVER, J. (1952) A dictionary of psychology, Harmondsworth: Penguin

EGGIESTON, J.F. (1973) 'Action and reaction in science teaching' inaugural lecture, Nottingham Univ. reported in Wragg (1975).

EGGIESTON, J.F., GALTON, M.J. & JONES, M. (1975) 'A conceptual map for interaction studies' in Chanan & Delamont (1975).

EVANS, K.M. (1952) A study of teaching ability at the training college stage in relation to the personality and attitudes of the student, Ph.D. thesis, Univ. of London.

EVANS, K.M. (1959) Research on teaching ability, Ed. Res. 1, 22-36.

EYSENCK, H.J. (1947) Dimensions of personality, London: Routledge and Kegan Paul.

EYSENCK, H.J. (1952) The scientific study of personality, London: Routledge and Kegan Paul.

FLANDERS, N.A. (1965) 'Teacher influence, pupil attitudes and achievement', Co-op. Research Monogr. No. 12 (OE-25040), US Office of Educ.

FLANDERS, N.A. (1970) Analyzing Teaching Behavior, Reading, Mass: Addison-Wesley.

FLANDERS, N.A. (1973) 'Knowledge about teacher effectiveness' - paper presented to the American Educational Research Assoc. New Orleans, Report 1973-17, Far West Laboratory for Educational Research and Development.

FOSTER, J. (1971) A factorial study of creativity in junior school children, M.Ed. thesis, University of Manchester.

FRICK, J.W. (1959) A factor-analytic study of flexibility in thinking, Educ. Psychol. Meas. 19, 469-96.

FRUCHTER, B. (1948) The nature of verbal fluency, Educ. Psychol. Meas. 8, 33-47.

GARNER, J. (1973) The nature of teaching and the effectiveness of teachers - In Lomax, D.E. (Ed.) The education of teachers, London: Wiley.

GARNETT, J.C.M. (1919) General ability, cleverness and purpose, B.J.Educ. Psychol, 9.

GETZELS, J.W. and
JACKSON, P.W. (1962) Creativity and intelligence: exploration with gifted students, London: Wiley.

GETZELS, J.W. & JACKSON, P.W. (1963) The teacher's personality characteristics - In Gage, N. (Ed.) Handbook of research on teaching, New York: Rand McNally.

GOLDMAN, R.J. (1964) The Minnesota Tests of Creative Thinking, Educ. Res. 7, 3-14.

GOUGH, H.G., DURFLINGER, G.W. & HILL, R.E. (1968) Predicting performance in student teaching from the California Psychological Inventory, J. Educ. Psychol. LIX, 119-127.

GUILFORD, J.P. (1950) Creativity, American Psychologist, 5, 444-454

GUILFORD, J.P. (1956) The structure of intellect, Psychol. Bull., 53, 267-293.

GUILFORD, J.P. (1959) Personality, New York: McGraw-Hill

GUILFORD, J.P. (1959a) Three faces of intellect, American Psychologist, 14, 469-479.

GUILFORD, J.P. (1959b) Traits of creativity, in Vernon, P.E. (Ed.) 1970.

GUILFORD, J.P. (1965) Fundamental statistics in psychology and education, New York, McGraw-Hill.

GUILFORD, J.P. (1967) The nature of human abilities, New York: McGraw-Hill.

GUILFORD, J.P., FRICK, J.W. et al (1956) A factor analytic study of flexibility in thinking - Reported in Guilford (1959).

GUILFORD, J.P. & CHRISTIANSEN, P.R. (1956) A factor analytic study of verbal fluency - Reported in Guilford (1959).

GUILFORD, J.P. and
MERRIFIELD, P.R. (1960 The structure of intellect model: its uses and implications, Rep. Psychol. Lab. no. 24, Los Angeles: University of Southern California, quoted in Torrance, E.P. (1962).

HADAMARD, J.S. (1945) An essay on the psychology of invention in the mathematical field, Princeton, New Jersey, Princeton University Press.

HAEFELE, J.W. (1962) Creativity and innovation, New York: Reinhold

HALL, C.S. & LINDZEY, G. (1970) Theories of personality, New York: Wiley.

HALL, E. (1970) School placement and teaching marks, Education for Teaching, Summer 1970, 55-57.

HANDLEY, G. (1973) Personality, learning and teaching, London: Routledge and Kegan Paul.

HARGREAVES, D.H. (1972) Interpersonal Relations and Education, Routledge and Kegan Paul.

HARGREAVES, H.L. (1927) The "faculty" of imagination, B.J. Psychol. Monograph Supplement No. 10.

HARTSHORNE, H. & MAY, M. (1928) Studies in the nature of character, Vol.1.
HARTSHORNE, H. & MAY, M. (1929) Studies in the nature of character, Vol.2, Reported in Wright, Taylor et al (1970).

HARVEY, O.J., HOFFMEISTER, J.K. et al (1971) A partial evaluation of Torrance's Tests of Creativity, Amer. Educ. Res. J. 7, 359-372.

HASAN, P. & BUTCHER, H.J. (1966) Creativity and Intelligence... B. J. Educ. Psychol. 57, 129-135.

HENJUM, A. (1969) A study of the significance of student teachers' personality characteristics, J. Teacher Educ. 20, (2).

HILGARD, E.R., ATKINSON, R.C. & R.L. (1971) Introduction to psychology, New York: Harcourt, Brace & Jovanovich.

HOUGH, J.B., LOHMANN, E.E. & OBER, R. (1969) 'Shaping and predicting verbal teaching behaviour in a general methods course' J. Teacher Ed., 20, 213-224.

HOUSTON, J.P. and MEDNICK, S.A. (1963) Creativity and the need for novelty, J. Abnorm. Soc. Psychol. 66, 137-144.

HOYLE, E. (1969) The role of the teacher, London: Routledge and Kegan Paul.

HUDSON, L. (1966) Contrary imaginations, Methuen.

HUGHES, D.C. (1973) 'An experimental investigation of the effects of pupil responding and teacher reacting on pupil achievement, Amer. Ed. Res. J. 10, (1), 21-37.

JACKSON, M. (1974) Inquiry into Personality Testing, Times Educational Supplement, 14th June.

JACKSON, P.W. and MESSICK, S. (1965) The person, the product and the response: conceptual problems in the assessment of creativity, - in Cashdan, A. and Whitehead, J. (Eds.) 1971.

JOYCE, B. & HODGES, R. (1966) Instructional flexibility training, J. Teacher Educ. XVII, 409-416.

JUNG, C.G. (1933) Psychological types, New York: Harcourt, Brace and World.

KLEINBERG, S.M. (1975) 'Interaction analysis and teacher training', Brit. J. Ed. Studies, 1, (3), 359-362.

KNOELL, D.M. (1953) The prediction of teaching success from word fluency data, J. Educ. Res. 46, 673-683.

KOESTLER, A. (1964) The act of creation, Hutchinson.

KRETSCHMER, E. (1925) Physique and character, London: Routledge and Kegan Paul.

LADUKE, C.V. (1945). The measurement of teaching ability, Study No. 3, J. Exp. Educ. 14, 75-100.

LAMKE, T.A. (1951) Personality and teaching success, J. Exp. Educ. 19, 217-259.

LANTZ, D.L. (1967) The relationship of university supervisors' and supervising teachers' ratings to observed student teaching behaviour, Amer. Educ. Res. J. 4.

LOMAX, D.E. (1969) The characteristics of successful student teachers, Ph.D. thesis, Univ. of Manchester.

LOVELL, K. (1951) An investigation into factors underlying teaching ability in primary and secondary modern schools with a view to improving methods of selecting potential teachers, M.A. thesis, Univ. of London.

LYKKEN, D.G. (1971) Multiple factor analysis of personality research, J. Exp. Res. in Personality, 5, 161-70.

McCLAIN, E.W. (1968) 16 P.F. scores and success in student teaching, J. Teacher Educ. 19, 25-30.

MACLEOD, G.M., MACLENNAN, D. & McINTYRE, D. (1975) 'Teachers' questions and reactions: a microteaching study', in Chanan & Delamont (1975).

McKNIGHT, P.C. (1971) 'Microteaching in teacher training: a review of research', Research in Educ. 6, p. 24.

McNEMAR, Q. (1964) Lost : our intelligence? Why?, American Psychologist, 19, 871-882.

MAITZMAN, I. (1960) On the training of originality, Psychol. Rev. 67, 229-242.

MARCUSE, H. (1955) Eros and civilization, Boston (Mass.) Beacon Press.

MEDLEY, D. & MITZEL, H.E. (1963) 'Measuring classroom behaviour by systematic observation' in Gage, N.L. Handbook of Research on Teaching, Chicago: Rand McNally.

MEDNICK, S.A. (1962) The associative basis of the creative process, Psychol. Rev. 69, 220-232.

MONTROSS, H.W. (1954) Temperament and teaching success, J. Exper. Educ. 23, 73-97.

MORGAN, J.C. & WOERDEHOFF, F.J. (1969) Stability of student teacher behaviors and their relationship to personality and creativity factors, J. Educ. Res. 62, (6), 251-254.

MORRISON, A. (1973) 'The teaching of international affairs in Scotland' Journal of the Modern Studies Assoc. No. 2, May, 1973.

MORRISON, A. & McINTYRE, D. (1973) Teachers and Teaching, Harmondsworth: Penguin Books.

NORRIS, R. (1975) An examination of schedules of criteria related to teacher competence, B.J.Teach.Educ. 1, 87-95.

NUTHALL, G. & SNOOK, J. (1973) 'Contemporary models of teaching' - in Travers, R.M.W. (ed) The Second Handbook of Research on Teaching, Chicago: Rand McNally.

PANTON, T.H. (1932) The assessment of teaching ability with special reference to men students in training, M.A. thesis, Univ. of London.

PARKER, D.H. (1963) Schooling for individual excellence, Nelson.

PINSENT, A. (1933) Pre-college teaching experience and other factors in teaching success of university students. B. J. Educ. Psychol. 3, 109-126 and 201-221.

POPPLETON, P.K. (1975) 'The classroom setting as a unit of analysis in observation studies' in Chanan & Delamont (1975).

PRINCE, M. (1924) The unconscious - Reported in Allport (1961).

PRINGLE, M.L.K., BUTLER, N.R. and DAVIE, R. (1967) 11,000 Seven-year-olds. First Report of the National Child Development Study (1958 cohort). Longmans.

RETAN, G.A. (1943) Emotional instability and teaching success, J. Educ. Res. 37, 135-141.

ROBERTSON, J.D.C. (1957) An analysis of the views of supervisors on the attributes of successful graduate student teachers, B.J.Educ. Psychol. 27, (2).

ROLFE, J.F. (1945) The measurement of teaching ability, Study no. 2. J. Exp. Educ. 14, 52-74.

RORER, L.G. (1965) The great response style myth, Psychological Bulletin, 63, (3), pp. 129-156.

ROSENSHINE, B. & FURST, N.F. (1971) Research on Teacher Performance Criteria in Smith, B.O. (ed) Research in Teacher Education, Englewood Cliffs, N.J: Prentice-Hall.

ROSENSHINE, B. & FURST, N.F. (1973) The Use of Direct Observation to Study Teaching, in Travers, R.M.W. (ed) The Second Handbook of Research on Teaching, Chicago: Rand McNally.

ROSTKER, L.E. (1945) The measurement of teaching ability, Study no.1. J. Exp. Educ. 14, 6-51.

RYAN, K. (1966) The teaching intern: a sociological stranger, J. Teacher Educ. XVII, 185-191.

RYANS, (1961) Some relationships between pupil behaviour and certain teacher characteristics, J. Educ. Psychol. 52, 82-90.

SCHWARTZ, A.N. (1950) A study of the discriminating efficiency of certain tests of primary source personality traits of teachers, J. Exp. Educ. 19, 63-94.

SEMEONOFF, B. (1970) Personality assessment, Harmondsworth: Penguin Books.

SHIPMAN, M.D. (1967) Education and college culture, B.J.Sociol. 18, (4).

SIMON, A. & BOYER, E.G. (eds) (1970) Mirrors for Behaviour, Philadelphia: Research for Better Schools.

SINCLAIR, J. McH. & COULTHARD, R.N. (1974) Towards an Analysis of Discourse, Oxford University Press.

SKINNER, B.F. (1968) The Technology of Teaching, New York: Appleton-Century-Crofts.

SORENSON, G. (1967) What is learned in practice teaching? J. Teacher Educ. 2, 173-178.

SPEARMAN, C. (1929) Address to the World Congress of Psychology, quoted in Crocker, A.C. (1973).

SPRINTHALL, N. et al (1966) A study of teacher effectiveness, J. Teacher Educ. XVII, 93-106.

START, K.B. (1974) Establishing childrens' learning as a criterion for teacher effectiveness, Ed. Res. 16, (3).

STONES, E. (1976) Teaching teaching skills, B.J.Teach.Educ. 2, (1), 59-78.

STONES, E. & MORRIS, S. (1972a) Teaching practice: problems and perspectives, London: Methuen.

STONES, E. & MORRIS, S. (1972b) The assessment of practical teaching, Ed. Res. 14, 110-119.

STUBBS, M. (1975) 'Teaching and talking: a sociolinguistic approach to classroom interaction', in Chanan and Delamont (1975).

TARPEY, M.S. (1965) Personality factors in teacher trainee selection, B.J.Educ. Psychol. 35, 140-149.

TAYLOR, C.W. (1947) A factorial study of fluency in writing, Psychometrika, 12, (4), 239-262.

TAYLOR, C.W. (ed.) (1964) Creativity: progress and potential, New York: McGraw-Hill.

THORNDIKE, R.L. (1971) Educational Measurement, Washington D.C.: Amer. Council on Education.

THURSTONE, L.L. (1938) Primary Mental Abilities, Psychometric Monograph No. 1.

TORRANCE, E.P. (1962) Guiding creative talent, Englewood Cliffs, New Jersey: Prentice-Hall.

TORRANCE, E.P. (1963) Creativity and intelligence, Minnesota Bureau of Ed. Res. Spring Report, - quoted in Foster (1971).

TORRANCE, E.P. (1966a) Torrance Tests of Creative Thinking, directions manual and scording guide, Lexington, Mass. Personnel Press.

TORRANCE, E.P. (1966b) Torrance Tests of Creative Thinking, norms-technical manual, Princeton, New Jersey: Personnel Press.

TORRANCE, E.P., TAN, C.A. & ALLMAN, T. (1970) Verbal originality and teacher behavior: a predictive validity study, J. Teacher Educ. XXI, (3), 335-341.

VELDMAN, D.J. & PECK, R.F. (1963) Student teachers' characteristics from the pupils' viewpoint, J.Educ.Psychol. 54, 346-355.

VERNON, P.E. (1939) Educational abilities of training college students, B.J.Educ. Psychol. 9, 233-250.

VERNON, P.E. (1962) The structure of human abilities, London: Methuen.

VERNON, P.E. (1963) Personality assessment: a critical survey, London, Methuen.

VERNON, P.E. (Ed.) (1970) Creativity, Harmondsworth: Penguin Books

VINACKE, W.E. (1952) The psychology of thinking, New York: McGraw-Hill.

VON HADEN, H.I. (1946) An evaluation of certain types of personal data employed in the prediction of teaching efficiency, J. Exper. Educ. 15, 61-84.

WALLACH, M.A. & KOGAN, N. (1965) Modes of thinking in young children, New York: Holt, Rinehart & Winston.

WARBURTON, F.W., BUTCHER, H.J. & FORREST, G.M. (1963) Predicting student performance in a university dept. of education, B.J.Educ. Psychol. 33, 68-79.

WARD, L.B. & KIRK, S.A. (1942) Studies in the selection of students for a teachers' college, J. Educ. Res. 35, 665-672.

WARREN, H.C. (1934) Dictionary of Psychology, Boston (Mass.) Houghton-Mifflin.

WATSON, J.B. (1925) Behaviorism, New York: Norton.

WEBSTER, R. (1976) The assessment of P.G.C.E. courses, B.J.Teach.Educ. 2, (1), 79-86.

WERTHEIMER, M. (1945). Productive thinking, New York: Harper.

WILSON, R.C. et al (1954) A factor analytic study of creative-thinking abilities, Psychometrika, 19, 297-311.

WRAGG, E.C. (1972) An analysis of the verbal classroom interaction between graduate student teachers and children, Ph.D. thesis, Univ. of Exeter.

WRAGG, E.C. (1973) 'A study of student teachers in the classroom' in Chanan, G. (ed) Towards a Science of Teaching, Slough: N.F.E.R.

WRAGG, E.C., OATES, J. & GUMP, P.V. (1975) 'Classroom Interaction', Block 11 of Open University Course E201.

WRIGHT, C.J. & NUTHALL, G. (1970) 'Relationships between teacher behaviours and pupil achievement in three experimental elementary lessons', American Educational Research Journal, 7, (4), 477-91.

WRIGHT, D.S., TAYLOR, A. et al (1970) Introducing psychology: an experimental approach, Harmondsworth: Penguin Books.

YAMAMOTO, K. (1962) A study of the relationships between creative thinking abilities of fifth grade teachers and academic achievement - Reported in Goldman (1964).

YAMAMOTO, K. (1965) Effects of restriction of range and test unreliability, etc. B.J.Educ.Psychol. 35, 300-305.

YAMAMOTO, K. (1964) Experimental scoring manuals for Minnesota Tests of Creative Thinking and Writing, Kent, Ohio: Bureau of Educ. Research.

APPENDIX I

Tables of Data

16 P.F. Questionnaire (Forms A & B) raw, sten, extraversion and anxiety scores. N = 13 men

Personality Factors

	A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4	Ex.	Anx
SB	20	17	28	25	29	28	30	32	15	23	18	29	15	10	23	28		
	9	7	5	6	6	7	6	10	5	6	4	8	4	2	5	6	7.5	6.4
GD	25	21	35	28	44	26	39	19	13	22	18	9	9	14	28	25		
	8	10	7	6	9	6	8	6	4	5	4	3	1	3	6	6	9.3	3.7
KJ	16	15	21	33	44	14	13	11	18	21	12	17	13	33	20	34		
	5	6	3	8	9	2	6	3	6	5	2	5	3	10	4	8	6.9	7.1
AK	16	19	40	10	16	26	26	26	8	24	22	12	21	19	33	12		
	5	8	8	1	3	6	5	8	1	6	6	4	7	5	8	3	2.9	2.2
CMcC	23	16	31	33	46	23	46	18	18	24	12	28	14	11	15	28		
	7	6	6	8	10	5	9	6	6	6	2	8	4	2	2	6	10	6.4
PW	23	19	34	22	28	19	26	26	24	20	20	15	24	24	28	26		
	7	8	6	5	6	4	5	8	8	5	5	4	8	7	6	6	5.3	5.6
TB	22	17	41	20	30	17	27	28	19	13	25	26	14	4	16	20		
	7	7	8	4	6	3	5	9	6	2	7	7	4	1	3	5	6.2	5.9
BD	22	24	37	39	40	28	22	22	16	39	12	16	22	18	16	31		
	7	10	7	9	8	7	5	7	5	10	2	5	7	5	3	7	7.7	6.1
GH	25	14	24	18	36	16	31	18	17	18	17	9	19	21	32	28		
	8	5	4	4	7	3	6	6	5	4	4	3	6	6	8	6	6.3	4.5
NS	28	20	38	15	36	26	42	33	17	26	19	12	20	12	20	21		
	9	9	7	3	7	6	8	10	5	7	5	4	6	3	4	5	7.8	4.2
AW	20	14	38	30	43	16	51	18	25	30	17	24	5	12	27	22		
	6	5	7	7	9	3	10	6	8	8	4	7	1	3	6	5	10	4.9
PWa	22	16	28	39	40	20	42	24	12	26	12	22	22	15	19	32		
	7	6	5	9	8	4	8	8	3	7	2	6	7	4	4	7	9.4	5.6
TW	20	12	34	22	24	32	26	28	22	30	19	41	12	21	6	32		
	6	4	6	5	5	8	5	9	7	8	5	10	3	6	1	7	4.9	8.1

16 P.F. Questionnaire (Forms A & B) raw, sten, extraversion and anxiety scores. N = 12 women.

Personality Factors

	A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4	Ex.	Anx.
PB	19	20	28	21	36	5	15	26	15	23	14	34	21	22	10	34		
	4	9	5	6	8	1	4	6	5	6	2	8	8	6	1	7	5.5	7.2
GHa	12	15	34	26	29	20	34	23	15	22	25	23	17	22	31	27		
	2	6	7	8	6	4	7	5	5	6	7	5	6	6	8	5	6.4	3.9
VW	19	15	26	16	25	24	15	27	19	25	23	24	14	20	12	34		
	4	6	5	5	5	5	4	6	7	7	6	8	5	6	2	7	4.0	8.2
LR	20	21	29	21	43	22	19	25	14	23	16	30	22	12	9	36		
	5	10	6	6	9	4	5	5	5	6	3	7	8	3	1	7	7.2	7.3
YJ	19	20	38	23	37	24	37	22	9	28	15	17	19	18	28	26		
	4	9	8	6	8	5	8	9	3	8	2	3	7	5	7	5	8.0	2.7
GHo	19	19	31	24	36	26	36	31	13	22	14	16	16	9	18	24		
	4	8	6	7	8	6	8	8	4	6	2	3	5	2	4	4	7.2	4.3
IP	19	17	27	5	21	17	15	22	8	19	20	32	15	24	16	28		
	4	7	5	1	4	3	4	4	2	4	4	7	4	7	3	5	2.2	5.9
JB	20	16	17	16	21	31	2	21	22	16	13	50	12	26	15	49		
	5	7	2	5	4	8	1	4	8	4	2	10	4	8	3	10	1.9	10
AC	11	17	25	24	28	23	16	23	8	40	10	26	15	31	24	31		
	2	7	5	7	6	5	4	5	2	10	1	6	5	10	6	6	3.8	5.3
RD	19	17	26	14	20	24	24	37	8	39	30	22	18	24	22	27		
	4	7	5	4	4	5	6	9	2	10	9	5	6	7	5	5	4.1	4.5
KW	20	12	16	13	30	26	20	28	20	34	8	35	20	23	24	36		
	5	4	2	4	6	6	5	7	7	9	1	8	7	7	6	7	4.6	7.8
EW	18	22	32	18	33	26	33	33	14	34	21	16	26	12	12	34		
	4	10	7	5	7	6	7	9	5	9	5	3	9	3	2	7	6.9	5.3

APPENDIX II

Torrance Tests of Creative Thinking (Forms A & B) Mean Scores. (Verbal scales)

Men	Fluency	Flexibility	Originality
SB	116.5	45	82
GD	89.5	48	45
KJ	100	49	52
AK	125.5	55.5	84.5
CmC	130	52	70.5
PW	47	27	32
TB	88	47	64
BD	76	46	51
GH	53	34	34
NS	92	44	54
AW	56	38	46
PWa	50	37	34
TW	49	25	31
Women			
PB	124	59	85
GHa	114.5	52.5	71
VW	111	59	73
LR	123	53.5	84.5
YJ	112	57	79
GHo	87	43.5	73
LP	125	48.5	71
JB	83	43	40
AC	84	47	53
RD	53	29	36
KW	82	44	46
EW	89	47	51

APPENDIX III

F.I.A.C. TALLY SHEET

Student:.....School.....

Class.....Subj/Activity.....

Date.....Time.....

FIAC	TALLY	TOTAL	PERCENT
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
	TOTAL		

Remarks

APPENDIX III a

Men

Student	Subject Affiliation	School	Lesson observed	Age and composition of class
SB	Physical Education	Hey Green Junior	Mathematics (3)	9/10 mixed
GD	Physical Education	Shorefield Comp.	Geography (2)	11/12 mixed
KJ	Physical Education	Bankfield Comp.	Geography (2)	11/12 mixed
AK	Divinity	S.Michael's Junior	History (3)	8/9 mixed
CMcC	Geography	West Bank Junior	Topic (pets) (2)	8/9 mixed
PW	Physical Education	Quarry Bank Comp.	Geography (2)	11/12 mixed
TB	History	Withernsfield Middle	History (2)	11/12 mixed
BD	Art & Craft	Lingham Middle	Painting (2)	10/11 mixed
GH	French	Lingham Middle	Discussions after story (2)	10/11 mixed
NS	Geography	New Heys Comp.	Geography (2)	11/12 mixed
AW	English	Lingham Middle	Discussions before drama (2)	10/11 mixed
PWa	American Studies	All Saints C.E.Junior	Red Indians (2)	9/10 mixed
TW	Art & Craft	West Bank Junior	Mathematics (3)	10/11 mixed

Women

Student	Subject Affiliation	School	Lessons observed	Age and composition of class
PB	English	Fairfield Junior	Mathematics (2)	10/11 mixed
GHa	History	Dovedale Junior	Mathematics (3)	9/10 mixed
VW	History	Malvern Junior	Project (2) (Communication)	9/10 mixed
LR	Geography	Fairfield Junior	Religion (3)	8/9 mixed
YJ	Art & Craft	Quarry Mt. Middle	Painting (2)	12/13 mixed
GHo	Physical Education	Holt Comp.	Biology (3)	11/12 mixed
LP	History	S.Michael's Junior	Mathematics (3)	9/10 mixed
JB	History	Farnworth C.E.Junior	Painting (3)	8/9 mixed
AC	Music	West Bank Junior	Discussions (3) after stories	8/9 mixed
RD	Divinity	Farnworth C.E.Junior	Project (2) (Undersea)	8/9 mixed
KW	Divinity	All Saints C.E.Junior	Mathematics (2)	9/10 mixed
EW	English	Farnworth C.E.Junior	Project (3) (Circus)	9/10 mixed

APPENDIX III b

Flanders' Interaction Analysis

Totals of tallies registered in each category with percentage of all tallies

Men	Categories										
	1	2	3	4	5	6	7	8	9	10	Total
SB	0	103	15	190	293	54	39	133	9	85	921
	0	11.2	1.6	20.6	31.8	5.9	4.2	14.4	1.0	9.2	100%
GD	1	58	8	92	261	24	23	92	32	34	625
	0.2	9.3	1.3	14.7	41.8	3.8	3.7	14.7	5.1	5.4	100%
KJ	0	63	0	148	127	45	0	91	3	24	501
	0	12.7	0	29.5	25.3	9.0	0	18.1	0.6	4.8	100%
AK	2	39	9	142	576	144	34	96	7	63	1112
	0.2	3.5	0.8	12.8	51.8	12.9	3.1	8.6	0.6	5.7	100%
CMcC	0	30	4	68	178	80	30	45	25	20	480
	0	6.3	0.8	14.2	37.1	16.7	6.3	9.4	5.2	4.2	100%
PW	1	57	8	83	211	79	35	56	34	60	624
	0.2	9.1	1.3	13.3	33.8	12.7	5.6	9.0	5.4	9.6	100%
TB	1	68	12	132	196	36	1	116	32	80	674
	0.2	10.1	1.8	19.5	29.0	5.3	0.2	17.2	4.7	11.8	100%
BD	0	48	6	78	150	132	45	60	66	42	627
	0	7.7	1.0	12.4	23.9	21.1	7.2	9.6	10.5	6.7	100%
GH	0	37	4	70	154	70	22	49	16	46	468
	0	7.9	0.9	15.0	32.9	15.0	4.7	10.5	3.4	9.8	100%
NS	1	65	17	105	197	21	11	73	5	33	528
	0.2	12.3	3.2	19.9	37.3	4.0	2.1	13.8	0.9	6.3	100%
AW	2	19	10	58	124	49	10	85	64	46	467
	0.4	4.1	2.1	12.4	26.6	10.5	2.1	18.2	13.7	9.9	100%
PWa	0	23	17	27	311	23	57	17	35	60	570
	0	4.0	3.0	4.7	54.6	4.0	10.0	3.0	6.1	10.5	100%
TW	1	61	10	166	217	43	43	109	67	116	833
	0.1	7.3	1.2	19.9	26.1	5.2	5.2	13.1	8.0	13.9	100%

Flanders' Interaction Analysis

Totals of tallies registered in each category with percentage of all tallies.

Women	Categories										
	1	2	3	4	5	6	7	8	9	10	Total
PB	0	15	5	25	275	70	20	25	10	20	464
	0	3.2	1.1	5.4	59.1	15.1	4.3	5.4	2.2	4.3	100%
GHa	0	7	0	70	84	223	63	70	42	56	615
	0	1.1	0	11.4	13.7	36.3	10.2	11.4	6.8	9.1	100%
VW	0	42	6	51	179	111	54	45	14	47	549
	0	7.7	1.1	9.3	32.6	20.2	9.8	8.2	2.6	8.6	100%
LR	0	15	3	54	508	33	45	33	21	36	748
	0	2.0	0.4	7.2	67.9	4.4	6.0	4.4	2.8	4.8	100%
YJ	3	74	9	95	226	143	12	63	21	21	667
	0.4	11.1	1.3	14.2	33.9	21.4	1.8	9.4	3.1	3.1	100%
GHo	0	69	23	92	207	103	92	138	46	46	816
	0	8.5	2.8	11.3	25.4	12.6	11.3	16.9	5.6	5.6	100%
LP	4	48	4	140	319	139	74	72	21	49	870
	0.5	5.5	0.5	16.1	36.7	16.0	8.5	8.3	2.4	5.6	100%
JB	3	106	13	117	217	48	30	115	88	60	797
	0.4	13.3	1.6	14.7	27.2	6.0	3.8	14.4	11.0	7.5	100%
AC	0	86	5	143	188	164	53	119	8	83	849
	0	10.1	0.6	16.8	22.1	19.3	6.2	14.0	0.9	9.8	100%
RD	0	40	4	76	238	78	13	78	2	45	574
	0	7.0	0.7	13.2	41.5	13.6	2.3	13.6	0.4	7.8	100%
KW	0	55	36	61	172	61	19	69	2	24	499
	0	11.0	7.2	12.1	34.5	12.1	3.8	13.8	0.4	4.8	100%
EW	1	67	13	115	145	16	22	121	40	43	583
	0.2	11.5	2.2	19.7	24.9	2.7	3.8	20.8	6.9	7.4	100%

APPENDIX IIIc

Flanders' Interaction Analysis; derived measures of percentage Teacher Talk, Teacher Question Ratio, Teacher Response Ratio and Pupil Initiation Ratio.

Men	TT%	TRR	TQR	PIR
SB	75.3	55.9	39.3	6.5
GD	74.8	59.0	26.0	25.8
KJ	76.5	58.5	53.5	3.2
AK	85.1	22.0	19.8	6.5
CMcC	81.4	30.9	27.7	35.6
PW	76.0	36.7	28.2	37.5
TB	66.1	68.8	40.2	21.5
BD	73.1	23.5	34.2	52.2
GH	76.4	36.7	31.3	24.5
NS	79.0	72.0	34.8	6.1
AW	58.2	34.4	31.8	42.9
PWa	80.3	33.3	7.9	67.0
TW	65.0	45.3	43.3	37.9
Women				
PB	88.2	19.1	8.4	28.9
GHa	72.7	2.3	45.4	37.4
VW	80.7	22.7	22.2	24.1
LR	87.9	18.8	9.6	38.9
YJ	84.1	35.6	29.7	24.8
GHo	71.9	32.1	30.8	24.9
LP	83.8	21.0	52.3	22.4
JB	67.0	61.0	35.1	56.7
AC	75.1	29.6	43.2	6.0
RD	78.3	32.6	24.1	2.9
KW	80.9	53.2	26.1	2.8
EW	65.0	68.1	44.2	24.9

APPENDIX IV

Analysis of F.I.A.C. Tallies for Different Occasions of Observation.

Men	Categories											Grand
	1	2	3	4	5	6	7	8	9	10	Total	Total
SB	0	39	6	65	102	17	14	48	2	29	322	
	0	14	2	21	38	9	6	22	2	14	128	
	0	50	7	104	153	28	19	63	5	42	471	921
GD	1	23	3	33	82	6	9	30	15	11	213	
	0	35	5	59	179	18	14	62	17	23	412	625
KJ	0	29	0	67	61	19	0	36	1	11	224	
	0	34	0	81	66	26	0	55	2	13	277	501
AK	0	12	2	39	179	47	11	32	2	19	343	
	2	14	5	71	249	65	14	30	4	37	481	
	0	13	2	32	148	32	9	34	1	17	288	1112
CMcC	0	21	2	41	102	41	18	23	17	13	278	
	0	9	2	27	76	39	12	22	8	7	202	480
PW	1	16	3	21	61	24	19	18	10	17	190	
	0	41	5	62	150	55	16	38	24	43	434	624
TB	0	22	5	42	81	14	1	46	15	29	255	
	1	46	7	90	115	22	0	70	17	51	419	674
BD	0	31	3	43	91	77	26	31	40	25	367	
	0	17	3	35	59	55	19	29	26	17	260	627
GH	0	17	2	29	69	34	12	24	7	24	218	
	0	20	2	41	85	36	10	25	9	22	250	468
NS	0	32	9	49	87	9	7	29	3	14	239	
	1	33	8	56	110	12	4	44	2	19	289	528
AW	0	7	3	17	35	16	6	24	19	17	144	
	2	12	7	41	89	33	4	61	45	29	323	467
PWa	0	10	9	12	139	9	31	11	17	26	264	
	0	13	8	15	172	14	26	6	18	34	306	570
TW	0	11	1	27	55	7	10	20	11	18	160	
	1	34	6	96	25	25	25	60	31	67	470	
	0	16	3	43	37	11	8	29	25	31	203	833

Analysis of F.I.A.C. tallies for Different occasions of Observation

Women	Categories 1	2	3	4	5	6	7	8	9	10	Total	Grand Total
PB	0	8	3	14	155	39	11	16	6	12	264	
	0	7	2	11	119	31	9	9	4	8	200	464
GHa	0	3	0	32	41	103	29	34	17	26	285	
	0	2	0	22	28	73	21	19	13	16	194	
	0	2	0	16	15	47	13	17	12	14	136	615
VW	0	7	0	11	31	26	12	9	2	10	108	
	0	35	6	40	148	85	42	36	12	37	441	549
LR	0	6	1	22	201	15	19	12	9	12	297	
	0	7	2	27	236	15	20	17	10	19	353	
	0	2	0	5	71	3	6	4	2	5	98	748
YJ	2	49	6	52	146	98	7	41	14	16	431	
	1	25	3	43	80	45	5	22	7	5	236	667
GHo	0	31	13	38	101	52	41	55	28	20	379	
	0	24	7	30	66	33	29	47	8	17	261	
	0	14	3	24	40	18	22	36	10	9	176	816
LP	1	15	2	37	83	32	20	20	6	11	227	
	1	9	0	32	78	40	15	18	5	10	208	
	2	24	2	71	158	67	39	34	10	28	435	870
JB	1	29	4	35	60	14	8	32	28	17	228	
	1	21	2	26	43	9	7	24	17	14	164	
	1	56	7	56	114	25	15	59	43	29	405	797
AC	0	8	1	14	16	18	6	12	0	6	81	
	0	22	1	38	51	44	14	29	3	23	225	
	0	56	3	91	121	102	33	78	5	54	543	849
RD	0	22	3	41	129	51	9	46	2	27	330	
	0	18	1	35	109	27	4	32	0	18	244	574
KW	0	16	14	20	64	22	6	27	0	8	177	
	0	39	22	41	108	39	13	42	2	16	322	499
EW	0	20	4	23	39	3	6	31	12	16	154	
	0	18	2	29	43	5	4	37	12	11	161	
	1	29	7	63	63	8	12	53	16	16	268	583

APPENDIX V

Product-moment correlations between FIAC and derived measures, and personality and creativity scores.

Men N = 13

	Personality/Creativity factor																			
FIAC	A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4	Ex.	Anx.	Flu.	Flex.	Orig.
1	-13	03	70	-40	-16	-14	32	15	35	-05	58	-11	-39	-29	32	-68	-03	-45	-18	-08	-04
2	41	17	-26	-12	12	-04	-26	13	16	-34	10	-12	-03	12	-17	23	-04	29	18	06	04
3	48	07	34	-04	03	08	47	60	00	09	15	04	11	-58	-09	-19	35	-19	-30	-22	-17
4	-07	-22	-29	-23	-03	-07	-28	-07	22	-29	13	15	-32	30	-25	15	-25	39	32	15	21
5	-01	16	06	-16	-18	10	25	17	-72	-17	03	-33	37	-07	37	-37	-02	-62	18	24	15
6	-23	25	-03	18	10	-09	-23	-39	06	34	-36	-22	30	25	15	04	23	-04	08	13	09
7	19	12	-17	02	08	-32	04	-08	-21	44	-50	-11	47	-01	17	42	07	12	-32	-33	-33
8	-04	-20	09	-16	15	-22	07	-15	43	-21	29	13	-74	-08	-08	-11	-08	16	15	14	15
9	-15	-07	32	49	30	-02	24	-22	52	57	-21	30	-34	-17	-25	18	40	22	-56	-39	-45
10	06	-51	10	-14	04	00	-21	40	43	-08	39	55	-06	-18	-30	13	35	37	-69	-76	-45
TT	07	35	-28	-13	30	13	-09	03	-72	46	-24	-48	64	31	28	-10	-17	-38	49	42	32
TRR	40	-03	-03	-25	09	-06	05	29	12	-47	35	03	-30	-24	-21	-02	08	14	13	07	04
TQR	-08	-14	-20	-07	02	-07	34	-12	46	-09	07	28	-39	26	-37	27	-19	56	11	-02	-06
PIR	-06	-19	06	64	30	-01	20	-14	20	46	-42	23	11	-10	-29	39	40	24	-63	-45	-57

size of r required to be significant at .05 level = .55

at .01 level = .68

Product-moment correlations between FIAC and derived measures, and personality and creativity scores.

Women N = 12

	Personality/Creativity Factor																				
FIAC	A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4	Ex.	Anx.	Flu.	Flex.	Orig.
1	27	16	09	-46	-26	15	-12	-18	-16	-39	-06	16	-22	10	02	14	-26	-03	15	06	-04
2	19	-12	-36	-17	-26	67	-10	22	18	29	-40	08	-09	12	09	33	-22	14	-61	-32	-61
3	36	-42	-52	-22	08	30	04	20	45	22	-53	11	25	-11	10	18	02	24	-34	-22	-34
4	-27	02	-01	-25	-42	51	14	11	-31	36	03	-25	-03	18	29	-03	-23	-27	-44	-43	-59
5	55	38	-06	-17	42	-49	-30	05	-04	-13	-08	21	38	-21	-60	16	12	29	35	20	41
6	-64	-42	37	32	-16	-28	30	-21	-20	-05	26	-21	-36	34	70	-52	07	-47	29	30	27
7	-31	-21	16	14	-02	-08	11	-38	07	-43	13	-09	-41	-22	-04	-22	-01	-05	36	24	40
8	-18	-10	-11	05	-24	61	31	37	04	39	-01	-33	08	-12	22	-01	-01	-20	-72	-60	-69
9	01	14	03	20	-07	37	-03	-28	46	-58	02	19	-18	-23	-11	44	-01	20	-08	01	-12
10	-67	-35	-18	10	-56	27	-18	-21	03	26	37	05	-39	41	16	08	-40	06	-38	-27	-47
TT	29	05	09	-16	31	-62	-15	-07	-24	-09	-10	11	-17	04	-16	-21	10	02	-57	43	61
TRR	37	05	-41	-26	-14	64	-11	29	32	25	-31	11	21	-09	-16	50	-13	28	-60	-40	-66
TQR	-50	-19	09	-19	-49	30	16	-23	-28	03	12	-19	-28	25	43	-20	-29	-32	-13	-23	-34
PIR	15	24	09	20	15	04	-23	-50	46	-83	-02	39	-17	-24	-33	50	03	35	34	36	29

size of r required to be significant at .05 level = .58

at .01 level = .71

APPENDIX VI

The Multiple Stepwise Regression Analyses

Summary Table 1.

Data derived for men

Criterion		Variables in order selected up to max. mult.R
F.I.A. Category	1	C, G, Q1, Q3, Flex, Q2, I, H, Q4, M
	2	A, Q2, M, Q3, Flex, N, G. E. B. H.
	3	I, H, Flu, B, G. N, A, M, C, Q4
	4	Anx, E, Flx, O, Q1, Q2, I, B
	5	L, H, Flex, Q4, Q1, M, E, Flu, Orig, F
	6	I, Q1, M, Orig, L, N, A, O
	7	N, Q1, Flex, L, C, Orig, M, B, I, G, F
	8	Q1, N, Q2, F, Orig, I, Flu, Q3, E, Q4
	9	M, L, Q2, I, Flu, F, E, C, Flex, Q3
	10	Flex, Q2, Ext, B, E, Flu, I, H
Teacher talk		L, Q1, Flu, Orig, Q3, O, Ext, B, F
Teacher response ratio		M, Q3, A, Q1, Q2, I, F, Ext, L, G
Teacher question ratio		Anx, E, O, Q1, Q4, M, L, Q3, Ext, Flu
Pupil initiation ratio		E, Flu, Q4, Flex, L, Q3, I, Q1, Q2, M, N

Summary Table 2

Date derived for women

Criterion		Variables in order selected up to max. mult.R
F.I.A. Category	1	E, H, I, B, Q2, A, Flex, G, Q4
	2	G, N, Orig, I, F, Flex, B, Ext
	3	N, I, B, Q1, M, Orig, Q2, E, A
	4	Orig, Anx, Flu, G, Q3, N, F, E, A
	5	Q3, G, A, F, Q2, N, Ext, Flex, Q4
	6	Q3, Flu, N, Q1, G, Flex, A, B, Q2
	7	M, A, Q4, Anx, G, O, Orig
	8	Flu, H, Orig, Q3, I, Flex, E, Q2, A
	9	M, G, E, Orig, B, Flu, I, F, Q2, N
	10	A, F, Q3, G, N, B, Flu, E
Teacher talk		G,A, Orig, M. O, Q4, L, Ext, E
Teacher response ratio		Orig, A, F, Flex, C, N, Q3, H
Teacher question ratio		A, E, G, H, I, B, Ext, Q4, F
Pupil initiation ratio		M, Q4, C, Flu, A, Flex, E, B

Table 1 (i) F.I.A. Category 1

Variable	beta coefficient	r	percentage criterion variance
C	1.0548	.70	73.8
Q1	-0.5197	-.39	20.3
Q2	0.5638	-.29	(-) 16.4
Q4	0.2022	-.68	(-) 13.8
Q3	0.3772	.32	12.1
H	0.3219	.32	10.3
I	0.5005	.15	7.5
G	-0.3608	-.14	5.1
Flex	-0.0661	-.08	0.5
M	-0.0604	-.05	0.3

Table 1 (ii) F.I.A. Category 2

Variable	beta coefficient	r	percentage criterion variance
A	1.6522	.41	67.8
Q2	1.7814	.42	21.4
N	1.7450	.10	17.5
M	0.4221	-.34	(-) 14.4
Q3	-0.8243	-.17	14.0
E	0.6862	-.12	(-) 8.2
Flex	1.1193	06	6.7
H	0.1533	-.26	(-) 4.0
B	-0.2126	.17	(-) 3.6
G	-0.6294	-.04	2.5

Table 1 (iii) F.I.A. Category 3

Variable	beta coefficient	r	percentage criterion variance
I	1.1394	.60	68.4
H	0.7827	.47	36.8
Flu	-0.4901	-.30	14.7
A	-0.3050	.48	(-) 14.6
C	-0.1512	.34	(-) 5.1
B	0.5018	.07	3.5
Q4	0.1380	-.19	(-) 2.6
G	-0.2671	.08	(-) 0.6
M	-0.1455	.09	(-) 1.3

Table 1 (iv) F.I.A. Category 4

Variable	beta coefficient	r	percentage criterion variance
Anx	1.4972	.39	58.4
Q1	-0.6874	-.32	22.0
O	-1.1086	.15	(-) 16.6
E	-0.5517	-.23	12.7
I	0.6457	-.07	12.3
Flex	0.6697	.15	10.1
B	-0.2366	-.22	(-) 4.5

Table 1 (v) F.I.A. Category 5

Variable	beta coefficient	r	percentage criterion variance
L	-0.7643	-.72	55.0
Q4	-0.5928	-.37	21.9
H	0.6661	.25	16.7
Flu	0.8789	.18	15.8
Orig	-0.9384	.15	(-) 14.1
E	0.7906	-.16	(-) 12.7
Q1	0.2536	.37	9.4
F	-0.7455	-.10	7.5
M	-0.2904	-.17	4.9

Table 1 (vi) F.I.A. Category 6

Variable	beta coefficient	r	percentage criterion variance
I	-2.4268	-.39	94.7
N	1.4041	-.36	(-) 50.6
Q1	1.4401	.30	43.2
M	1.0495	.34	35.7
A	0.8541	-.23	(-) 19.6
O	0.6741	-.22	(-) 14.8
Orig	0.6290	.09	5.7
L	0.2310	.06	1.4

Table 1 (vii) F.I.A. Category 7

Variable	beta coefficient	r	percentage criterion variance
Flex	-2.7710	-.33	91.4
N	-1.4453	-.50	72.3
Orig	1.5039	-.33	(-)49.6
M	-0.7223	.44	(-)31.8
L	-1.0718	-.21	22.5
C	1.3201	-.17	(-)22.5
G	0.2765	.32	8.9
Q1	0.1855	.47	8.7
I	-0.5735	.08	(-) 4.6
F	0.4858	.08	3.9
B	0.0658	.12	0.8

Table 1(viii) F.I.A. Category 8

Variable	beta coefficient	r	percentage criterion variance
Q1	-0.5345	-.74	39.6
N	1.1389	.29	33.0
Orig	1.8901	.15	28.4
F	1.6085	.15	24.1
Flu	-1.3036	.15	(-)19.6
Q2	1.0793	-.08	(-) 8.6
I	0.3948	-.15	(-) 5.9
Q3	-0.3179	-.08	2.5
E	-0.3573	-.16	5.7
Q4	0.2153	-.11	(-) 2.4

Table 1 (ix) F.I.A.Category 9

Variable	beta coefficient	r	percentage criterion variance
M	0.4369	.57	24.9
E	0.3409	.49	16.7
I	-0.7404	-.22	16.3
L	0.2361	.52	12.3
Flu	-0.2187	-.56	12.3
F	-0.3801	.30	(-)11.4
Flex	-0.2875	-.39	11.2
Q2	-0.5893	-.17	10.0
C	0.3053	.32	9.8
Q3	0.0642	-.25	(-) 1.6

Table 1 (x) F.I.A.Category 10

Variable	beta coefficient	r	percentage criterion variance
Flu	-0.6894	-.69	47.6
B	-0.5767	-.51	29.4
I	0.2851	.40	11.4
Q2	-0.4692	-.18	8.5
H	-0.3319	-.21	7.0
Ext	-0.1979	.35	(-) 6.9
Flex	0.0813	-.76	(-) 6.2
E	0.2091	-.14	2.8

Table 1 (xi) Teacher Talk

Variable	beta coefficient	r	percentage criterion variance
Flu	1.5934	.49	78.1
Q1	0.7790	.64	49.9
Orig	-1.2487	.32	(-)40.0
L	-0.2650	-.72	19.1
Q3	0.3666	.28	10.3
O	0.1963	-.48	(-) 9.4
B	-0.1623	.35	(-) 5.7
F	-0.1820	.30	(-) 5.5
Ext	0.2638	-.17	(-) 4.5

Table 1 (xii) Teacher Response Ratio

Variable	beta coefficient	r	percentage criterion variance
I	1.8758	.29	54.4
Q1	-0.9967	-.30	29.9
M	-0.4075	-.47	19.2
F	2.0059	.09	18.1
Q2	0.5274	-.24	(-)12.7
Ext	-1.2630	.08	(-)10.1
L	-0.3397	.12	(-) 4.1
Q3	-0.1691	-.21	3.6
A	0.0507	.40	2.0

Table 1 (xiii) Teacher Question Ratio

Variable	beta coefficient	r	percentage criterion variance
Anx	3.7715	.56	211.2
O	-1.6247	.28	(-) 45.5
Q4	-1.3551	.27	(-) 36.6
L	-0.7331	.46	(-) 33.7
Q1	-0.6788	-.39	26.5
Q3	0.4282	-.37	(-) 15.8
E	-0.9920	-.07	6.9
Ext	0.2885	-.19	(-) 5.5
M	0.5534	-.09	(-) 5.0
Flu	-0.1456	.11	(-) 1.6

Table 1 (xiv) Pupil Initiation Ratio

Variable	beta coefficient	r	percentage criterion variance
Flex	-1.4974	-.45	67.4
E	0.9913	.64	63.4
Q4	-0.8046	.39	(-) 31.5
Flu	0.3378	-.63	(-) 21.3
I	-0.8363	-.14	12.1
Q3	-0.3528	-.29	10.2
N	0.2409	-.42	(-) 10.1
M	0.2186	.46	10.1
L	-0.4466	.20	(-) 8.9
Q2	-0.5091	-.10	5.1
Q1	0.4377	.11	4.8

Table 2 (i) F.I.A. Category 1

Variable	beta coefficient	r	percentage criterion variance
E	-0.7754	-.46	35.7
A	0.5408	.27	14.6
B	0.7948	.16	12.8
I	-0.6458	-.18	11.6
H	0.9637	-.12	(-)11.6
Q2	1.0199	.10	10.2
G	0.4008	.15	6.0
Q4	-0.2397	.14	(-) 3.4
Flex	0.3642	.06	2.2

Table 2 (ii) F.I.A. Category 2

Variable	beta coefficient	r	percentage criterion variance
N	-0.9256	-.40	37.0
G	0.4623	.67	31.0
F	-1.1505	-.26	30.0
Orig	-0.3273	-.61	20.0
Flex	0.5692	-.32	(-)18.2
I	0.5319	.22	11.7
Ext	0.2915	-.22	(-) 6.4
B	0.4284	-.12	(-) 5.1

Table 2 (iii) F.I.A. Category 3

Variable	beta coefficient	r	percentage criterion variance
N	-0.7970	-.53	42.2
B	-0.8030	-.42	33.7
I	0.6085	.20	12.2
Q1	0.4402	.25	11.0
M	-0.3251	.22	(-) 7.2
Orig	-0.2001	-.34	6.8
A	-0.1813	.36	(-) 6.5
E	-0.2662	-.22	5.9
Q2	-0.3083	-.11	3.4

Table 2 (iv) F.I.A. Category 4

Variable	beta coefficient	r		percentage criterion variance
Orig	-2.1220	-.59		125.2
Flu	1.8157	-.44	(-)	79.9
Anx	-1.4597	-.27		39.4
F	-0.7649	-.42		32.1
Q3	-0.8944	.29	(-)	25.9
G	0.5029	.51		25.7
E	0.4535	-.25	(-)	19.1
A	0.1724	-.27	(-)	4.7
N	-0.6250	.03	(-)	1.9

Table 2 (v) F.I.A.Category 5

Variable	beta coefficient	r		percentage criterion variance
F	2.5052	.42		105.2
A	0.6966	.55		38.3
Q2	0.9073	-.21	(-)	19.1
Ext	-1.3280	.12	(-)	15.9
Q3	-0.1212	-.60		7.3
N	0.8666	-.08	(-)	6.9
Flex	-0.2386	.20	(-)	4.8
Q4	-0.2098	.16	(-)	3.4
G	0.0400	-.49	(-)	2.0

Table 2 (vi) F.I.A.Category 6

Variable	beta coefficient	r		percentage criterion variance
Q3	0.6286	.70		44.0
A	-0.2264	-.64		14.5
Flex	0.4769	.30		14.3
G	-0.4051	-.28		11.3
Q1	-0.3084	-.36		11.1
B	-0.2425	-.42		10.2
Q2	-0.2454	.34	(-)	8.3
N	0.3122	.26		8.1
Flu	-0.1545	.29	(-)	4.5

Table 2 (vii) F.I.A. Category 7

Variable	beta coefficient	r	percentage criterion variance
M	-0.8319	-.43	35.8
Q4	-1.5757	-.22	34.7
A	-0.7258	-.31	22.5
Anx	2.0932	-.05	(-) 10.5
O	-0.8535	-.09	7.7
Orig	-0.0805	.40	(-) 3.2

Table 2 (viii) F.I.A.Category 8

Variable	beta coefficient	r	percentage criterion variance
Flu	-2.1051	-.72	151.6
H	2.8076	.31	87.0
I	-1.5596	.37	(-) 57.7
Flex	0.7444	-.60	(-) 44.7
Q3	-1.3810	.22	(-) 30.4
Q2	1.0071	-.12	(-) 12.1
Orig	-0.1750	-.69	12.0
A	0.3571	-.18	(-) 6.4

Table 2 (ix) F.I.A.Category 9

Variable	beta coefficient	r	percentage criterion variance
M	-0.9115	-.58	52.9
Orig	-1.8625	-.12	22.4
E	0.9338	.20	18.7
Q2	-0.4975	-.23	11.4
Flu	1.2954	-.08	(-) 10.4
I	0.2474	-.28	(-) 6.9
F	-0.7162	-.07	5.0
B	0.3142	.14	4.5
G	0.0579	.37	2.1
N	-0.2702	.02	(-) 0.5

Table 2 (x) F.I.A. Category 10

Variable	beta coefficient	r	percentage criterion variance
A	-0.8647	-.67	57.9
F	-0.2743	-.56	15.4
N	0.3039	.37	11.2
B	-0.3086	-.35	10.8
Q3	-0.6521	.16	(-)10.4
G	0.2958	.27	8.0
Flu	-0.1450	-.38	5.5
E	0.1873	.10	1.9

Table 2 (xi) Teacher Talk

Variable	beta coefficient	r	percentage criterion variance
Orig	0.5689	.61	34.7
O	2.1842	.11	24.0
Q4	-0.8731	-.21	18.3
L	-0.5071	-.24	12.2
Ext	1.0393	.10	10.4
E	-0.2730	.31	(-) 8.5
M	0.6871	-.09	(-) 6.2
A	0.1858	.29	5.4
G	0.0204	-.62	1.3
N	0.1027	-.10	(-) 1.0

Table 2 (xii) Teacher Response Ratio

Variable	beta coefficient	r	percentage criterion variance
Orig	-1.5740	-.66	103.0
C	0.5546	-.41	(-)22.7
Flex	0.5486	-.40	(-)21.9
N	-0.6255	-.31	19.4
A	0.4125	.37	15.3
Q3	-0.4418	-.16	7.1
H	0.3584	-.11	(-) 3.9
F	-0.1708	-.14	2.4

Table 2 (xiii) Teacher Question Ratio

Variable	beta coefficient	r	percentage criterion variance
Ext	-2.0294	-.29	58.9
H	2.2314	.16	35.7
F	0.3973	-.49	(-) 19.5
A	-0.3744	-.50	18.7
I	-0.4939	-.23	11.4
Q4	0.4779	-.20	(-) 9.6
B	0.4031	-.19	(-) 7.7
E	-0.2978	-.19	5.7
G	0.1758	.30	5.3

Table 2 (xiv) Pupil Initiation Ratio

Variable	beta coefficient	r	percentage criterion variance
M	-0.6987	-.83	58.0
Q4	0.8743	.50	43.7
Flex	-0.3177	.36	(-) 11.4
C	0.7067	.09	6.4
E	0.1438	.20	2.9
Flu	0.0697	.34	2.4
B	-0.0673	.24	(-) 1.6
A	-0.0897	.15	(-) 1.3